Consciousness and Freedom

THREE VIEWS

Pratima Bowes

Lecturer in Philosophy, University of Sussex

METHUEN & CO LTD

11 NEW FETTER LANE · LONDON EC4

First published in 1971 by
Methuen & Co Ltd
11 New Fetter Lane London EC4

Printed in Great Britain by
Ebenezer Baylis & Son Ltd
The Trinity Press
Worcester, and London

SBN 416 08720 5

Distributed in the U.S.A. by
Barnes & Noble Inc.

Contents

Introduction[1]

It is perhaps true to say that at the commonsense level most men take themselves to be possessed of a body as well as a mind which are believed to interact to produce a unitary personality. At this level this interaction is also believed to be a matter of experience. For instance, a consciously felt intention to close a door – and this is taken as a mental fact – actually leads to one's performing some movements which result in the door getting shut – and this is taken to be a physical activity. Equally, movement in the stomach, which is bodily activity, produces a conscious desire to eat, which is thought to be mental. How a physical activity, which can be measured in terms of physical and chemical laws, can lead to a mental experience which cannot be so measured, or vice versa, is a problem with which commonsense is not especially bothered. It is content to accept that this is the case. (Commonsense attitude in this respect is like that of Aristotle – what does happen can happen.)

Furthermore, at the commonsense level no clear analysis of what is meant by 'mind' is made. A desire to eat is mental but this 'mental' desire may be approached from two different levels: (1) desire to eat may solely consist of a propensity to search for food and perform certain intelligent and appropriate movements necessary to satisfy desire; (2) it may also appear as a conscious experience of felt awareness that one is having a desire to eat. This last, however, need not necessarily lead to activities involved in (1), for someone who is engaged in slimming may decide not to eat despite the desire. Anyway, these two facets of desire as a mental fact are not ordinarily separated, and as they often occur together in human experience they are taken to be facts of the same order. Both are called 'mental' or 'conscious' and either of these facets may be produced as an evidence for the occurrence of a

[1] This introduction was written at the suggestion of Professor Patrick Corbett who very kindly read the manuscript and thought that this should be added. I am extremely grateful to him for this suggestion for it enabled me to bring out, on a much clearer level, certain basic themes which occur throughout the book without being systematically put together.

fact which may be said to be 'mental'. For instance, when a baby cries at meal time, it is perfectly in order to say that he is feeling a desire for food, although he may have no conscious awareness that he is feeling it.

There is, however, a scope for making a distinction between the two aspects of mind. A dog too feels a desire to eat, that is, a propensity which may be said to be mental, without however that involving a conscious experience of felt awareness that it is feeling this desire.[1] Thus an animal mind is similar to a human mind in certain respects, but at the same time different in other respects, because it lacks a certain dimension necessarily possessed by a developed human mind (a degree of development achieved by all human minds during the process of growth from infancy onwards). It seems to me that it is because of this particular dimension of the human mind, its capacity to have a conscious awareness that it is having an experience, and not merely because of the dimension which is constituted of dispositional and performatory propensities of an intelligent kind that normally go with conscious experience, that a human being feels that he has a mind which is additional to his physical attributes. But as I said, no clear-cut distinctions between propensities to behave intelligently and conscious awareness of having an experience involving these propensities is made at the commonsense level, since both these dimensions of mind constitute integral parts of the human mode of existence and they are very much mixed up in what is called 'experience'.

Commonsense is content to leave all this vague and undefined, but philosophy is not. In the modern period, Descartes undertook a philosophical analysis of what the mind consists of and his finding was that mind is essentially consciousness. This is presented in his famous 'cogito' principle, 'I think therefore I exist'. All behaviour, according to him, is bodily and it is conducted purely mechanically except for the fact that in a human personality consciousness does somehow interact with the physical processes of the body and influence its functionings. Exactly how it does that or what precisely is involved in this influencing when physical activities are said to be mechanical processes, Descartes did not make very clear. His philosophy is thus refined commonsense dualism where refinement consists in demonstrating by means of philosophical analysis (whether or not one accepts this demonstration is another matter) that consciousness is a fact of its own kind which vindicates itself even when one tries to doubt its existence away, unlike the body, the existence of which can be doubted without contradiction. This to Descartes' way of thinking proves that the mind cannot be derived from the body. Descartes thus means by 'mind' consciousness of private experience (I think) and not merely experience, the evidence for which lies in performance or propensities to behave intelligently in relation to stimulus.

[1] Some people may object that we do not know that the dog does not have this as well. It seems to me that if the dog did have this capacity, it would have symbolically codified its experience in language as human beings do and used some such term as 'I'. It does not seem to me that a conscious awareness of private experience is possible without the use of 'I' or a similar term.

This of course leads to the difficult problem of interaction, that is, how can a conscious process which is not mechanical influence a physical process which is. This is supposed to contradict a basic principle of science that the total amount of energy is constant and activities of a physical state are causally linked with antecedent physical states. This would be upset if a conscious state which is supposed not to be a physical state can induce expenditure of energy involving a physical state in the body. Because of this, scientifically minded philosophers believe that they must question that consciousness acts on the physical processes at all.

This denial takes several forms: (1) Although consciousness is a fact of its own kind since it obviously cannot be confused with a physical state, whilst one is actually having a conscious experience it is just a by-product of physical processes in the brain and has no causal efficacy. But even if consciousness has no causal efficacy, its very presence as a fact of its own kind is an embarrassment to a materialist monist who believes that the only kind of entity that could possibly exist is the kind studied by science, that is, physico-chemical energy. So this gradually leads to the second kind of denial.

(2) Consciousness is not a fact of its own kind at all, but is actually identical with a brain state representing dispositional properties for intelligent performance. So it requires no causal agencies other than the physical and chemical ones involved in the occurrence of the brain state of which the conscious dimension is simply a characteristic logical expression, without being a different sort of entity. Now this scheme involves some difficulty. It is not impossible to persuade oneself to believe that a conscious state is causally produced by a physical state, but it becomes a little more difficult to believe that a state of conscious experience, involving an awareness that so and so, is literally identical with a state of expenditure of physico-chemical energy in the brain. But this identity must be posited if consciousness is not to be reckoned

as a thing of its own kind – the only kind of thing acceptable to science being physico-chemical energy. This difficulty as far as I can see is not solved by the identity theorists who, having proclaimed the identity thesis, go on to study mind purely in terms of dispositional and behavioural properties, thereby quietly ignoring the existence of conscious awareness of private experience. The dimension of mind which consists in dispositional and behavioural properties can be very profitably studied in neuro-physiological terms and the success achieved in this field has produced an illusion that the identity thesis has been experimentally vindicated. What the success of neurophysiology proves, however, is the identity with physical states of that particular dimension of mind which consists in dispositional and behavioural properties, not that of the dimension which consists in conscious experience of awareness that so and so – for this has not been even included in the terms of one's study.

(3) The third form of denial seems to arise out of an unconscious awareness that the literal identity of private experience with expenditure of physico-chemical energy in the brain is somewhat implausible, so it boldly asserts that there is no such thing as a state of private experience at all to be reckoned with. This it seeks to show by adopting the policy of scheme (2), that is, by taking mind as identical with that particular dimension of it which consists in dispositional and behavioural properties of an intelligent nature. 'Seeing red' consists in being able to discriminate red from green and so on and not in having a private show of one's own in the internal screen of the mind to be called 'the phenomenal field'. Consciousness is simply behaving, verbally or otherwise. Thus consciousness is gradually explained away rather than explained, and the sole candidate that remains on the stage to qualify as mind is that particular dimension of human (and animal) behaviour which consists in performance, from simple stimulus-response, like stopping when the red light is on, to logical operations of a complex kind like mathematical deduction.

What is studied in this scheme is called 'mind' and legitimately so, because performance of this kind has a mental dimension. But although it leaves aside conscious experience of private awareness as irrelevant to our understanding of human behaviour, it nevertheless goes on to identify the dimension of the mental it does study with consciousness and claim that a sufficient explanation of consciousness has been given in neurophysiological terms. It thus follows commonsense in equating mind with consciousness but with a fundamental difference. In the commonsense scheme this equation leads to a belief that conscious experience has something to do with how we behave; in the behaviourist-materialist scheme it leads to the claim that since mind in the sense of disposition for intelligent behaviour is nothing other than a property of a physical state, consciousness, if it exists, is nothing other than this either.

The behaviourist-materialist scheme of studying man claims to be scientific and so it is, insofar as its method of studying neurophysiological correlations involved in intelligent behaviour is concerned. It is perfectly justified in claiming that intelligent behaviour it is studying can be explained without the hypothesis of the influence of consciousness on physical states if it finds that no such influence is needed to explain the facts it is studying. But is it scientific to claim either that consciousness is not a fact of its own kind or that it never influences human behaviour in any way at all? The practice of science requires the methodological assumption that the phenomena it is studying are explainable in terms of correlations between different expenditures of physico-chemical energy. It is not a part of this assumption to claim that nothing but physico-chemical energy exists in the world. If one makes this assumption one is adopting a scientific world view, and this is metaphysics, not science. As metaphysics, it is just as much a matter of value judgment as the contrary metaphysics which denies that matter can exist at all without the support of mind. The fact that this metaphysics is advanced in the name of science makes no difference to its status as metaphysics.

Science as such does not claim to study all that there can possibly be in the universe of experience, it only claims to explore the nature of the physical world. From this it follows that whatever can be studied successfully by the method of science is physical, but it does not follow that whatever cannot be or is not being studied by science does not exist, or that if it does *appear* to exist it must be an illusion, nor even that it must be one day, with further progress in science, accessible to scientific methodology. All of these assumptions are concealed value judgments inspired no doubt by a faith in science, but not forced on us by the existence of science or its methodology.

Behaviourism of all kinds, by adopting the method of science in its study of human beings, has decided to study those aspects of human experience and behaviour that are amenable to scientific methodology and hence are physical in nature, even though aspects studied represent highly complex intelligent performance like logical deduction or

purposive behaviour of various degrees of organisation. That there is no difficulty in conceiving of intelligent and purposive behaviour, which may be called 'mental', as a product of matter, is shown by the adoption of precisely this view by the Sāmkhya school of philosophy in India, which is a very ancient school of thought and is thus not influenced by the practice of science. But unlike behaviourism Sāmkhya claims that a human personality exhibits mind as well as consciousness and this latter cannot be explained by matter. Anyway, as I said, neurophysiology or behaviourism as a discipline has decided that an explanation of human behaviour is to be approached from the point of view of finding correlations between neural states and physical activities and this already shows a preference, for it is not impossible for there to be a radically different scheme of explanation suited to some other purpose. And since as science it can only study what its methodology allows it to study, it leaves aside that aspect of human behaviour and experience which is not amenable to the method it uses in its approach to the problem of understanding how the human mind functions.

Behaviourism does not of course recognise this and claims instead that this other aspect does not exist. But what it is expressing in saying that consciousness does not exist or it does not influence behaviour is a decision to study that particular aspect of mind which consists in intelligent performance because this lends itself to scientific methodology and is thus of interest to science. The aspect that does not lend itself to scientific study is of no interest to the behaviourist and therefore might as well not exist for the purpose of science.

That fact is that if consciousness exists and it does influence human behaviour, this will not be caught by a methodology which is designed on the assumption that it does not exist or that it does not influence human behaviour, because the facts it will be studying will be precisely those facts which are free of the hypothesis of consciousness and its supposed influence. The trouble is that untutored human experience suggests that consciousness does exist and that it does influence behaviour.[1] So the adoption of a methodology that leaves this out already expresses a value judgment that the type of facts it studies is the only type that is important for us to understand about man, particularly if this adoption is supplemented by a denial that whatever is

[1] This does not require us to suppose that there can be any expenditure of physical energy which is not already under-written by antecedent physical states. This will be discussed in the first chapter.

not required as a hypothesis for the purposes of this methodology does not in actual fact exist. It is another matter if neurophysiology simply says that it is studying those aspects of human behaviour and experience which can be explained purely physically in accordance with the method of science, leaving it open as to whether there exist aspects which cannot be so explained.

It seems to me that on its own ground neurophysiology is a valuable study which can enormously extend our understanding of how a human being functions, even if this is not the only kind of understanding we can possibly have of man. The trouble is that value judgments which are made in the name of science appear to be purely scientific and therefore value-neutral. But this is very deceptive and prevents us from realising that a scientific world view is not the same thing as science. Neurophysiology as a branch of science which studies human behaviour from a particular angle of approach is one thing – and there may be no end of things we can find out about behaviour by studying it in this manner – but neurophysiology as a complete philosophy of mind is quite another. For in advancing this philosophy it is going beyond what it can legitimately talk about in its capacity as pure science, that is, about its data consisting of correlations found between physical and mental states (mental meaning intelligent and purposive behaviour) and the assumption – involved in finding these data – that an explanation can be found for intelligent and purposive behaviour in terms of physico-chemical, electrical activities of brain states. Both the data and the assumption can be perfectly legitimate without its being true that nothing else need ever be involved in our understanding of human behaviour, no matter what the purpose for which we want to understand it. That one goes on to assume this last as well, even though it is not strictly necessary for the success of neurophysiology as a scientific enterprise, shows that one is involved in a value judgment and not merely in stating accredited facts. This judgment is concealed under scientific terminology, but it is nevertheless there, and this explains the appeal of this philosophy to people of a certain mental orientation who feel that we cannot be said to have knowledge of anything unless it can be studied in accordance with the method of science despite the fact that it obviously leaves out things that we intuitively know about ourselves. It is possible that it is precisely because our acceptance of the reality of consciousness and its influence on behaviour depend upon our intuitive understanding of ourselves as a being of a certain kind – a self-conscious being who is present to

oneself and to whom the world is present – that scientifically minded philosophers feel a prejudice against it. Strangely enough, prejudice against intuition is far more deeply rooted in scientific philosophy than in science itself, although this prejudice is justified in the name of science and its methodology. There is increasing recognition in the field of science[1] that no scientific facts are discovered purely by using a methodology and that methodology in science serves to establish on acceptable grounds – through processes of verification, falsification, etc. – what one discovers by intuition. No doubt intuition can go wrong just as perception and inference can, but it is there and is crucial in our understanding of facts, even in the domain of science. Of course, what one intuitively discovers in science can be subsequently subjected to the requirements of scientific methodology. This is not possible with consciousness, so prejudice against it is hard to overcome.

But it is not only materialists or behaviourists whose philosophy of mind is inspired by a value judgment. Some thinkers who very much acknowledge the reality of consciousness and its role in behaviour seem to suffer from it equally. These thinkers are not content to recognise that consciousness is there and that it has an important role in human existence; they make consciousness the central fact of human existence in terms of which all behaviour is to be understood. What it amounts to, however, is a decision to study those aspects of behaviour in which consciousness appears to play a central role because this is what really interests the thinkers concerned – and not the aspect which can be explained as an intelligent reaction to a given stimulus. But as this is not consciously recognised, it is felt that consciousness must be claimed to be central to all behaviour. Both Husserl and Sartre labour under this view of consciousness which constitutes the whole of the mind as far as they are concerned. As a result any evidence that a piece of behaviour is mental, that is to say, discriminative, intelligent, purposive, etc., is taken as a sufficient ground for claiming it is inspired, constituted or at least attended by consciousness. This, of course, is as much contrary to our experience of ourselves as the materialist doctrine that consciousness does not exist or that it does not have any causal role. So this insistence on the centrality of consciousness in face of contrary evidence really represents a value judgment that what is important in our understanding of a human being is the fact that he is self-conscious and is capable of operating in a manner in which

[1] See P. B. Medawar, *Induction and Intuition in Scientific Thought* (Methuen, London, 1969).

nothing that is not of this nature can operate – other things being unimportant or uninteresting.

According to Husserl human beings live in an inter-personal cultural world which can only be shared by using concepts, and the meaning of these concepts can only be understood when consciousness constitutes the logical structure of these concepts by intuitively grasping their essence. Man not only intuitively knows himself to be conscious, he also intuitively understands the various logical structures involved in the use of concepts in inter-personal communications, and this he can do because he is possessed of consciousness which is capable of constituting this logic. This does not however appear to be the way in which we understand the meaning of concepts, for we seem to be able to use them meaningfully without even being consciously aware that we are actually using them, let alone intuitively grasping their essence. It seems to me that Husserl does not sufficiently distinguish between reflective understanding of experience, and this is not possible without consciousness, and understanding which is shown at the ordinary level of meaningful behaviour where consciousness need not even appear. What he is analysing in phenomenology is reflective understanding of concepts and at this level we do come to see that words have a logical structure – even if this logic is contextual and not fixed once and for all – which can be said to be constituted in consciousness and intuitively grasped on a self-conscious plane. But we can operate perfectly well without any reflective understanding of our performance, and if there is a logical structure involved in speech the mind can grasp this at the level of performance (use of speech, etc.) purely by virtue of innate intelligent capacities of brain states. But this aspect obviously does not interest Husserl.

Thus it is not true that we do not understand the meaning of experience that requires use of concepts without the active intervention of consciousness. But it is another matter with reflective understanding of experience, where meaning is grasped on a self-conscious plane. The fact that Husserl does not distinguish between these two kinds of understanding – this is shown by his failure to mention the presence of understanding at the level of performance where consciousness is not needed as a hypothesis – and that he puts forward reflective understanding as the model of all human understanding of experience, show that he is making a value judgment. This judgment has two facets, a judgment to the effect that reflective understanding of experience on a self-conscious plane is the most typical way of being human as well as

to the effect that this is the most important thing for us to understand about the human mode of existence. I agree with Husserl that the fact of consciousness is important in the human scheme of operation; also that possession of self-consciousness makes man, in some ways at least, a unique kind of being, without however rejecting the behaviourist thesis, which Husserl may be disposed to deny, that a great deal about human understanding of meaning of experience can be understood purely in terms of non-self-conscious processes of the mind.

Sartre goes to even greater lengths in claiming the centrality of consciousness. As an existentialist he is no doubt committed to the recognition of the facticity of man based on his embodiment and his situation. These things are mentioned but their practical implications are more or less ignored when Sartre embarks on an exploration of man's nature and his possibilities. Man is defined as the for-itself, which means consciousness, and contrasted with the in-itself, which means matter. One wonders how the fact of one's having a body and the various laws under which it functions fits in with this definition. For man can behave purposefully and intelligently by virtue of his having a body of a certain kind without consciousness being in any way involved. But this is not allowed by Sartre, possibly because this does not interest him. Simply because consciousness is conscious of itself we are said to be translucently aware of all our motivations and springs of action. And if we persuade ourselves that we do not know about these that is because we know in order not to know. Clearly this follows only if consciousness coincides with the very being of man, not if it constitutes only a certain dimension of his being. As consciousness coincides with the very being of man according to Sartre and consciousness cannot be derived from matter, man is not only self-conscious through and through (to deny this is to indulge in bad faith), he is freedom through and through as well. Now so far is this from being a true account of how human beings experience themselves that Sartre's assessment of man as a being of this nature can only be a value judgment – presented no doubt in factual terms – to the effect that man ought to strive to counteract his tendency to behave like an object and aim at being as self-conscious and free as he possibly can. The urgency of this value judgment is understandable in face of theories which express a contrary value judgment that consciousness if it exists is quite unimportant and even irrelevant and that man is a thoroughly determined being. Even so, his doctrine about man remains a value judgment concerned to say what man ought to be like considering his

possibilities as a self-conscious being, under the guise of a factual assertion as to what he is like but does not want to be, as well as a decision to study man only insofar as he fits the requirement.

The Indian theories discussed in this book have quite a different objective from Sartre but equally with Sartre they find that the most exciting and promising thing about man is his consciousness. They agree with Sartre that consciousness reveals an object while revealing itself at the same time. But what to them is the most significant fact about consciousness is its capacity to function merely as a witness of all that goes on in one's psychophysical system – just watching almost like an outsider in spite of the fact that what is going on is part of oneself. This achievement of a distance with regard to one's own personality that is involved in the witnessing function of consciousness forms the basis of one's freedom from psychophysical determinism. Thus even though the empirical man is involved in the psychical and physical functions of his personality a part of him, his consciousness, can by taking up the attitude of a witness, remain uninvolved and thereby transcend the limitations of his psychophysical being as a determined system. This of course is freedom which characterises the very nature of consciousness as witness and being freedom it represents a state of perfect bliss or at least absence of all suffering to which an empirical man is subject.

I said that consciousness can take up the attitude of a witness (this of course is possible because of the very nature of consciousness as self-consciousness). But the aim of these Indian systems being to persuade men to liberate themselves from all material limitations, they prefer to say in a categorical manner that the true or real being of man consists in being this witness. From this it follows that it is a mistake for a man to think that he is his mind-body system at all. Here again this rejection of an obvious fact really amounts to a decision that man is an interesting being only insofar as he can take up the attitude of witness remaining quite uninvolved in all that may be happening to himself in order that he may experience himself as unlimited or identical with all existence. The fact that man has a mind and a body which limit him to the narrow compass of his individual existence is not an interesting fact about man from this point of view. Thus it is apparent that when they say that man is nothing but the witnessing function of his consciousness, what they are expressing is a value judgment that this is the most significant fact about man and that man ought to believe himself to be essentially this in order that aspects of his being which get in the

way of his realisation of freedom may be got rid of. But as I said, they prefer to make a factual statement to the effect that man is his witnessing function of consciousness, and a totally unconditioned being. This creates a contradiction, for we are told on the one hand that the mind-body system is just not a part of man, and yet advised on the other to make the utmost effort to get rid of what is not at all a part of ourselves.

An attempt is made in these systems to solve this contradiction by making a distinction between the transcendental man, the real or true man, and the empirical man subject to ignorance and mistake about his true being. But the point is that if the very nature of the empirical man is constituted out of ignorance or mistake then at the empirical level at any rate it remains true that man is his mind-body system whether or not he is consciousness as well and it is not a mistake for him to think that this is the case. But Sāmkhya and Vedānta, because they believe that what they are doing is stating simple facts about what is the case rather than preferences for certain facts at the cost of others, are forced to say – despite the distinction made between the transcendental and empirical man – that it is a mistake for the empirical man to believe that he is his mind-body system even while he remains solidly grounded in the empirical plane of experience. This implies that the empirical man ought not to be there at all, for his being empirical means his taking his mind-body system seriously as a part of himself. This paradoxical implication of their utterances could be avoided if Sāmkhya and Vedānta preached that man should behave *as if* his mind-body system is not an essential part of his being, and he can behave like this because of the capacity of his consciousness for transcendence in the manner of a witness, instead of stating in a descriptive vein that the mind-body system is not a part of man.

Thus it happens that a great deal of what is called philosophy of mind is concealed value judgment based on a decision to study mind insofar as it fulfils a certain requirement. A philosopher finds that one or the other dimension of mind is more important or valuable, depending on what he thinks human life is like or ought to be like, and on the basis of this preference builds a theory to show how what he takes mind to consist of – intelligent performance explainable purely on physico-chemical, electrical operations of brain states, consciousness that is conscious of itself or consciousness that can detach itself from all particular states and processes and thus experience itself as unlimited – is the central fact by means of which everything about man, his behaviour and experience, can be accounted for. I said preference, because an exclusive acceptance of one or the other dimension of mind is not forced upon us by the nature of things and people in widely different intellectual and cultural contexts and at different points of their historical development have come upon both these dimensions of what is called the 'mind' – the one that is determined by a given stimulus and explainable on the basis of the functioning of the body – and the other which is self-conscious and because of the capacity of consciousness for transcendence, free. That our understanding of mind as determined and physically based[1] is not necessarily a product of modern scientific attitude is shown by the fact that this was already grasped by an ancient philosophical system in India, although it used this understanding to an entirely different end than what is espoused by its modern protagonists.

Again, the understanding of human mind as self-conscious and in that capacity as something radically different from matter is not a doctrine that was introduced for the first time in human thinking by Descartes in the seventeenth century (who is often accused of having introduced a pseudo problem in philosophy—mind-body relation). For this too is an ancient Indian doctrine. It seems natural then that these two dimensions of mind can be found not only on a commonsense level where our understanding of what a human being is like may be vague and confused but also on the level of philosophical reflection where they can appear in a much more clear-cut fashion. The trouble is that these two dimensions when seen together suggest a view of man which is messy, ambiguous and even paradoxical: for they indicate at the same time that man is conscious and he can, in his capacity for transcendence inherent in his consciousness, influence his situation, both internal and external, to some degree at any rate, as well as that man's mental capacities and experiences are a function of the workings of his body. The indeterminate character of this situation may be all right for commonsense to put up with, for at this level man is content to accept a fact as a fact and leave it at that, but not for philosophers who feel uncomfortable – and this seems to be almost a built-in feature of their profession – unless they can find a unitary system which explains every conceivable fact within the domain of their investigation whether or not it appears to be amenable to this treatment. A fact which is not so amenable is usually characterised as a mere appearance having no basis in reality, an illusion fostered by

[1] The theory that consciousness is just a function of matter and is nothing apart from it was first propounded by Indian materialists as early as the fifth or sixth century B.C.

something or other, a distortion of the true state of affairs and so forth, all of which seem to mean that the fact concerned is unimportant and not worth bothering about. This applies to philosophy of mind just as much as to any other problem and we are often offered, at least in theory, a vision of man, who is either just brain and nervous system or nothing but pure consciousness, these two being difficult to put together in a unitary scheme.

Man however may be a complex being, and it is possible that a full understanding of how man does or can function requires that we have to approach the question from different levels or points of view, none of which can claim to possess the sole truth about him. This may be intellectually untidy but nearer to facts about man as we experience them. I am not however saying that philosophers ought not to have preferences or that preferences felt by them cannot be in any way justified. But this justification seems to me to lie in showing that one can hope to achieve some worthwhile result by acting or thinking in the line of these preferences rather than in simple statements of fact as to what is the case, when such statements obviously contradict aspects of our experience of ourselves and others. Behaviourist-materialist preferences can be justified in terms of our scientific understanding of human behaviour insofar as this is determined by brain states; phenomenological-existentialist preferences in terms of possibilities of man as a responsible being who can acknowledge his responsibility to create his own mode of existence insofar as this can be done by a self-conscious understanding on the part of man of his own performances; and Indian preferences in terms of achieving a transcendental dimension not disturbed by the ordinary limitations of the human condition insofar as this can be realised by the witnessing function of consciousness. All these objectives may be worthwhile, but none of them can claim to be exploring the only true or real nature of man, if man, as I believe to be the case, is a multidimensional being and it lies in his nature to be capable of functioning in all of these modes.

I

The Materialist-Behaviourist View

CONSCIOUSNESS

The materialist theory of mind – with which behaviourist theories are associated – may also be said to be a theory of no mind in some sense. Most people, when they talk about their minds, as distinct from their bodies, vaguely consider mind to be a distinctive kind of functioning which is not thought to be identical with the functioning of the body, although they certainly also believe that the functioning of the mind is intimately associated with the functioning of the brain. Even theories of mind which insist on the distinctness of mental functioning associate the occurrence of mental states with the occurrence of brain states whether the relationship between these two kinds of states is conceived under the model of parallelism or interactionism. The materialist theory denies that there are two distinct types of states at all to interact or to be parallel with one another – one occurring on the internal screen of the mind, as it were, the other in the publicly inspectable brain. There is just one kind of states, the brain states, with which the so-called mental states are identical.

Of course the materialist account is not that clear-cut. Sometimes materialist talk about mind is such that we could say that there are mental states but somehow these states are identical with brain states; at other times it looks as if the materialist is denying that there are any states occurring on their own in a psychical sphere which are at all identifiable as mental states. D. M. Armstrong in his *A Materialist Theory of the Mind* appears to hold this latter view and as this gives a plausible account of what are normally called 'mental states' I shall base my analysis mostly on this variety. But of course even Armstrong has to admit that there occurs consciousness, or that there is such a state as being conscious, which does not fit in very well with the

extend to all that pertains to it. It is this incomplete understanding of his own nature that sometimes creates in man fear about his own possibilities. Man is thus not necessarily acting in bad faith if he feels that he does not completely control his own springs of action or that he does not have any real choice.

3
Two Indian Views

I shall now consider some classical Indian views on the subjects of consciousness and freedom. The views I am going to consider here are not the only Indian views on these subjects. Indeed, classical Indians have held all kinds of opinion about consciousness beginning with the extreme materialistic claim of Cārvāka philosophy that consciousness is generated out of the mixture of four material elements, earth, water, air and fire, just as the intoxicating quality of alcohol is generated out of non-alcoholic matter, and ending with the extreme idealist claim of Vijnānavāda Buddhism that consciousness is the one reality and what is thought to be the material world simply does not exist. (This is to be differentiated from the somewhat different Samkarite view, to be discussed later, that the material world has an indescribable status, which is neither wholly real like consciousness, nor wholly unreal like a phantom.) The views I have chosen to discuss here, Sāmkhya and Samkara Vedānta, throw some light, so it seems to me, on the problems of consciousness and freedom and they may be thought of as further extensions of certain notions found in the phenomenological and existentialist doctrines I have discussed. (I do not mean of course that the phenomenologists and existentialists would agree with the Indian views.) I must make it clear here that I am not concerned with the schools of philosophy discussed in this chapter in their historical character, and as looked at in an orthodox manner, but shall freely interpret them in the light of our modern understanding of these problems without however imputing to them doctrines which cannot be thought to be strictly in accordance with their teachings as traditionally preached.

SĀMKHYA

One of the two philosophical schools I have chosen to discuss here is Sāmkhya, a very ancient school of thought – pre-Buddhist in its earliest formulation – which has of course gone through subsequent elaborations in the characteristic Indian manner through people writing commentaries, and further commentaries on these commentaries, on the main body of aphorisms embodying its philosophical position. The interest of this rather archaic school in the present context lies in this. It is a most uncompromising kind of dualism which divides the whole realm of experience into two kinds of being (or functionings) embodying two different principles, the principle of consciousness also called self, and the principle of materiality, also called nature, and says that there is no interaction between the two. It then goes on to make a distinction between consciousness or self – the being of which lies solely in its function of manifestation – and mind, conceived as intelligent activity based on a capacity (unconscious) for discrimination, sensory, intellectual and volitional. The whole realm of the mind is classified under the principle of materiality by Sāmkhya and not under consciousness or self. In saying that mental functions are material functions it is in agreement with that trend of modern thought which, impressed with purposeful, intelligent activities performed by machines, claims that the mental is nothing other than the physical. (Sāmkhya of course would say 'material', as the concept of the material is wider in Sāmkhya than the physical. This will be discussed later.) The difference lies in this of course, that having initially separated consciousness from the mental or the realm of the mind Sāmkhya is not constrained to say that man is nothing but a physical organism, since man has consciousness which a machine does not have even if we decide to say that the machine has a mind insofar as it has the capability to perform intelligently, purposefully and discriminatively.

This distinction between mind and consciousness is something that is foreign to Western philosophy and a Western philosopher thinks of consciousness as an attribute of the mental. This distinction, however, is vital to Indian philosophy and must be made clear before we can proceed. Of course there used to be a distinction made in Western thinking between mind and soul or spirit. In Indian thinking this soul or spirit in man is nothing other than the principle of consciousness that functions in him. It is not the case that the soul or spirit has the attribute of consciousness, rather what is called 'soul' or 'spirit' is nothing other than the function of consciousness as such. The principle of consciousness embodying a certain manner of being or functioning denotes a non-material and underived principle that functions in man and it is referred to as 'pure consciousness' or 'consciousness as such' to distinguish it from the empirical fact of an instance of conscious awareness of something on the part of a particular man at a particular time which is relative to the object known. We are not of course empirically aware of consciousness as such – understood in a topic-free way – in our normal experience of the world and many, including some classical Indian thinkers, would claim that there is no such thing as a topic-free consciousness at all. But this topic-free consciousness must not be thought to be an actual instance of conscious functioning, for there is no such instance so far as our empirical awareness of the world is concerned. It is the very principle of subjectivity as such which we have already encountered in Husserl and Sartre and, as I have emphasised earlier on, the concept of a principle of functioning and the concept of an instance of such functioning, each has a logic of its own not to be confused with the other.

Indian philosophers, when faced with the objection that there is no such thing as consciousness as such, meaning that there is no empirical experience of such a thing, stress that even if all consciousness is consciousness of something there must be a function called 'consciousness' to be conscious of this something. Many would object no doubt that this is hypostatising consciousness which arises only in a particular context of contact with objects and which is not to be thought of as an entity by itself, but the Indians claim that consciousness performs a distinct function, that of manifestation (equivalent to Sartre's revelation and Husserl's constitution function) of the object it is conscious of as well as of itself – a function which cannot be performed by anything which is non-conscious and so it must be thought of as there as a reality of a distinct sort. (Whether the function of consciousness is an entity or not is irrelevant to the question of its reality.) Furthermore they claim like Descartes that the being of consciousness establishes itself just through being itself, so it is a reality in respect of which doubt is not even possible. In the language of Sāmkhya it is self-contradictory to try to prove that it does not exist. No doubt all empirical consciousness is conscious of something and not of itself as a certain kind of function quite apart from this something. But as Husserl said consciousness can be understood as a certain kind of

function by means of a transcendental deduction. According to the Indians, it is also possible to experience consciousness just as a function of awareness, as distinct from the awareness of this or that, in a special kind of experience called transcendental self-experience which consists in consciousness's awareness of itself without any mediation of consciousness of objects. Husserl's transcendental deduction is a logical process that proceeds by an analysis of the very possibility of knowledge – Indian philosophers too use this sort of analysis without using this name of course; the Indian transcendental self-experience is supposed to be a direct experience of itself on the part of consciousness just as a function of manifestation and not as the specific manifestation of this or that object. Whether or not there can be such a transcendental self-experience is something which is subject to experiential verification; so, at any rate, is the claim.

Anyway, although empirically speaking all consciousness is consciousness of this or that and consciousness of this is quite distinct from consciousness of that there is no reason why we cannot use the concept of 'consciousness as such' to denote a certain kind of function of which both the consciousness of this and the consciousness of that, either on the part of subject X or subject Y, can be thought of as instances. We seem to have no difficulty in using the concept 'energy', which embodies the very principle of physical functioning, even though this principle is never encountered face to face on an empirical level where we find only instances of energy, chemical, electrical and so on. Consciousness, as a function of manifestation, is sometimes thought of in Indian philosophy as a kind of power or energy to be called 'spiritual' to demarcate it from the physical which, whatever tasks it may perform, does not perform that of self-manifestation. But whether or not we think of consciousness as power or energy it is vitally important for the Indian philosophers we are considering to grasp the concept of 'consciousness as such' (or the principle of consciousness embodying a manifestation function) as distinct from that of empirical consciousness which is an instance of awareness of this or that here and now. For their primary interest lies in exploring the transcendental feature of consciousness as a certain kind of principle of functioning rather than in any empirical functioning of this consciousness.

One of the reasons why people are inclined to feel that consciousness is a function of the brain is that they identify the conscious with the mental, and the mental, as recent researches in neurophysiology and computer functioning show, can be identified with the physical with some gain in clarity and understanding. If the mental is the physical the conscious must be physical too, for consciousness is an attribute that sometimes qualifies the mental. But the conclusion that the conscious is the physical does not follow if the conscious is something distinct from the mental. This is where Sāmkhya philosophy comes in, which may have a contribution to make, not in the details of its explanation, much of which is pretty archaic, but in its contention that the conscious is not the mental when mental is characterised by intelligence, and that the mental has to be explained in terms of the material. But before we go on to give a somewhat detailed account of what Sāmkhya has to say, let us see how we can meaningfully make a distinction between the mental and the conscious, which seem all of a piece, and what the term 'mental' is to mean when it is identified with the material.

The primary distinction between the mental and the purely physical – when a distinction between the two is meaningful – is that purely physical functioning is thought to be non-purposive, non-intelligent, non-discriminative. A piece of stone thrown with a certain force will move across space with a certain velocity and then fall somewhere, all according to certain laws of motion. The stone does not care whether it moved or not, whether it fell to the ground or where it fell, it has no purpose and its functioning is determined purely by certain antecedent incidents with which it has certain law-governed relationships. But if you push a man with the intention of throwing him on to the ground he, if not caught unawares, will resist your push and co-ordinate his muscle activities in order that he may do so successfully, varying his responses as the situation needs and without sticking to any given pattern of expenditure of force. The reactions of the man show capacities that are mental, that is, purposeful, discriminative and so on. But the man, particularly if he was not expecting you to push him, might not be consciously aware that he was being pushed or that he was doing things in order to resist that push, yet nevertheless he can deal with the situation perfectly intelligently without having to be so conscious and entirely on an unconscious plane of understanding of what is going on. That is, we can say that he perceived the situation as being so and so or was unconsciously aware of it – how otherwise could he deal with it – but not that he was *consciously aware that* the situation was so and so. Now if the man is acting intelligently but unconsciously we cannot say that he is being guided by his consciousness in doing what he is doing. How then does he manage to behave as he does?

His behaviour, as neurophysiology shows, can be explained by what

N

is happening in his brain and nervous system – nervous impulse being conducted along various channels, electrical discharge taking place at certain nerve cells etc. – processes and states which can all be studied strictly according to laws of physics and chemistry and the man's behaving as he does consists in nothing but the happening of some inner states and processes in the brain and the nervous system. As long as we confine ourselves to behaviour and response I do not see how we can reject this neurophysiological account as a complete explanation. Of course the behaviour is intelligent etc., but the neurophysiological states and processes and the mechanisms of their regulation are such that the resultant activity does show the characteristics of being intelligent and purposive. This is at first difficult to accept as we are accustomed to think that physical nature is blind, unintelligent and lacking in purpose. This is not so according to Sāmkhya which considers intelligence as a function of the material at a certain level of subtlety of composition. Anyway, the activities of computers must be recognised as intelligent and purposive, judging from their performance – and we use the same criterion of performance to judge if a man is acting intelligently. This shows that physico-chemical electrical operations can, with the introduction of feed-back mechanism, take on an intelligent, purposive and discriminative character and function on the level of self-correction, self-adjustment, self-regulation and so on. There is no reason why the intelligent activities of human beings cannot be explained on the same principles as the intelligent activities of machines.

What then can be the distinction between the mental and the physical when we can no longer say that the physical is necessarily non-intelligent, and why use two languages if they are the same? One way of marking this distinction is to say that the mental is the physical functioning at a certain level of complexity and organisation and anything below this level is merely physical. That is, physical activity, at a certain level of complexity and organisation, is purposive, intelligent, discriminative; below that it is blind, of an unvarying, uniform nature, lacking in discriminative perception. Another way is suggested by Armstrong. The same states which are considered instances of physical states functioning according to physio-chemical laws, may also be considered in their aspect as causal agencies, leading to or being the outcome of other states. In this latter capacity they are states apt for the production of certain behaviour and can be thought of as mental. It may then perhaps be said that the language which uses the term 'mental' sees a state in the perspective of a network of interrelated behaviour unlike the use of purely physical language. I think that this account can be slightly improved if we say that mental is that state which is apt for the production of discriminative behaviour, for it is only when action is discriminative that it is useful and necessary to think of it in the context of a network of interrelated behaviour. An unvarying blind response to a stimulus has not got this causal role in a system of actions as discriminative response has.

Where then does consciousness fit into this picture? As we have seen, it is not necessary for a man to be conscious in order to behave intelligently. Is a man then unaware of what he is reacting to when he behaves intelligently without being conscious that he is reacting in this way? I think that we should say that a man is aware but only unconsciously, which means that he is aware of whatever he is reacting to at the level of behaviour, insofar as he is reacting to it and his reactions are adequate to the situation, but not that he is consciously aware, that is, he is aware of whatever is the case but not aware that such and such is the case. Conscious awareness then represents knowledge that *p*. It may be objected that a man must know that something is the case, when he reacts to it. But here 'knowing that' is being used technically to demarcate conscious awareness from unconscious response which also involves awareness of some sort, and it cannot be said that these two types of awareness are not there to be distinguished. This is why in Indian philosophy the term 'consciousness' is often used interchangeably with 'knowledge'. Intelligent response is not knowledge that, it is only conscious manifestation that something is the case which represents knowledge that (of course in the vocabulary of Indian philosophy it is just referred to as 'knowledge' without the qualification that). But a state of conscious manifestation or knowledge that something is the case is also a case of self-manifestation, that is, it not only represents knowledge that something is the case it also represents knowledge that there is awareness that something is the case. Any state which does not involve this self-manifestation is not conscious according to the Indian use of the term 'conscious', however intelligent or discriminative may be the performance it leads to. Only that awareness is conscious which is conscious of itself and consciousness is thus defined as self-manifestation. This Indian idea of self-manifestation of consciousness closely corresponds to Sartre's pre-reflective cogito and here, just as in Sartre, the distinction between the conscious and the unconscious is a fundamental feature of reality. The conscious involves self-manifestation by

definition, and the unconscious lacks this capacity. These definitions however are not arbitrary, they are based on our experience of two distinct orders of functioning.

Here, as in Sartre, consciousness is not a state of one's mind, it is a function, the function of manifestation and more importantly, from the point of view of distinguishing it from unconscious awareness, of self-manifestation. To say that someone has become conscious that such and such is the case is not to say that he is going through a state, although he may be going through some state; it is to say that there has occurred an awareness on his part that such and such, of which fact he is aware. This function of self-manifestation is not an attribute of the mental, rather it manifests the mental as mental, in the same way as it manifests the physical as physical. So the distinction between the mental and the conscious is this, mental is a state apt for the production of discriminative behaviour, conscious is the awareness of this state as this state. Consciousness is often confused with what it manifests, so when it is said that all consciousness is consciousness of something it looks as if consciousness can be explained away in terms of this something. But if this something does not manifest itself as this something – meaning that it is not aware of itself as itself – there must be a manifesting function apart from this something, otherwise we cannot explain how the object gets manifested at all, that is, how the knowledge that it is such and such, comes about.

Thus the sole reality of consciousness lies in its manifestation function, both other and self-manifestation, for consciousness that such and such cannot take place unless such consciousness is conscious of itself, although awareness of p, as distinct from awareness that p, can take place on an unconscious plane. But such an unconscious awareness of p is not a manifestation of p which consists in the awareness that p. We shall see later how this manifesting function in its capacity as other-manifestation is limited, at the empirical level, by the mind and body of the person who is being conscious of whatever he is being conscious of. But the important thing is to see that there is no such thing, according to this model, as states of consciousness, or, if there is such a thing as a state of consciousness, it is the same state, a state of manifestation. What varies is the object manifested and the mental states which are modifications of these objects but not the function of manifestation as such. This is very similar to the Sartrean view that consciousness has no content, it is just a wind blowing towards the object. As there are no states there is no question of these states being generated by brain

activity, brain activity only decides what consciousness is going to manifest, it does not generate the manifesting capacity of consciousness itself. Indian philosophy does not of course talk of the brain, it talks of internal and external senses, the mind and external organs of sense. The manifesting capacity of consciousness as shown in the instances of its functioning in a human being is limited by its psychophysical states, for it can only manifest something which has produced some modification in a person via the activities of his internal and external senses. Consciousness as a pure function of manifestation is then like a searchlight which can be turned on various objects via the internal states of a person modified in accordance with the impact of these objects, but it turns on itself as soon as it is turned on something. There is nothing but consciousness which functions in this manner, so consciousness cannot be derived from anything but itself. As in Sartre it is self-existent.

It may be objected that one may be conscious of an object without being conscious of this fact. But here the term 'conscious' is not being used in the Indian sense. When one is being said to be conscious of an object without being conscious of the fact that there is this awareness, what is happening is that one is being unconsciously aware of the object, one does not know *that* the object, for if one knows this one also knows that one knows. The term consciousness is always used in this sense of self-manifestation in the Indian views we are considering. Of course when one knows that one knows x one does not necessarily formulate a proposition, 'I know that I know x'. Verbal response by itself whether overt or covert is not consciousness. It is, like every other human performance, a result of certain physical functions and consciousness manifests this response just as much as anything else. Anyway consciousness' self awareness is not always verbally articulate and Sartre's terminology 'pre-reflective cogito' expresses this. As he says, when one is conscious of an object one's consciousness of it is positional, it is directed towards this object, but implicit in this positional consciousness is also to be found a non-positional consciousness of itself, consciousness which is not consciously directed to itself as it is directed to this object, but nevertheless non-positionally aware of itself. In the Indian view consciousness does not need to know itself in the manner of an object, its very nature is self-illumination like that of light which illumines itself as well as objects just by being there. A so-called consciousness which is not conscious of itself will not be consciousness, it will be unconscious awareness. But there is in the world more than unconscious awareness, there is conscious awareness as well. And it is

this capacity for conscious awareness that distinguishes man from the rest of the world.

We may now have a brief look at Sāmkhya metaphysics. As already remarked, Sāmkhya divides the world into two opposite principles, the principle of consciousness also called self, but self is identical with the function of consciousness and is not something that has consciousness, – and the principle of materiality also called nature but not nature of solid objects or elements but of forces of functioning too subtle to be empirically perceptible. The principle of materiality is composed of three different strands of subtle force or energy and there is no specific manifestation or instance of the functioning of this principle so long as these three strands remain in perfect equilibrium. These are: (1) *Sattva* or forces responsible for physical manifestation such as that of light, for transparency such as that of crystal, and lightness as shown in the tendency of a flame to go upward (2) *Rajas*, or forces responsible for tension, drive and movement and (3) *Tamas*, forces representing heaviness and inertia which act as obstacles or restraining influences on the other two strands. It is because of this last strand that there is not a chaotic superabundance of material manifestation in nature wherein it introduces a lawful character by acting as a restraining force on strands of energy responsible for physical manifestation and movement. These strands have physical as well as psychological connotations both of which qualify for the description 'material'. The first makes for the psychological state of tranquillity and peace, the second restlessness and turmoil and the third stupidity, dullness and inertia.

It is not exactly clear on what grounds these strands of energy are chosen as the basis of all material functioning. Nor is it clear why the presence of consciousness and its interest in nature should disturb the equilibrium of these forces, a state which constitutes primordial matter in its infinite potentiality, thereby starting off the process of material manifestation. It is not impossible that Sāmkhya philosophers vaguely felt that it is only in the presence of consciousness that classification according to definite categories can take place and it is only when there is the possibility of such classification that there is manifestation. But this is perhaps too sophisticated an interpretation of a very ancient doctrine, although if I were writing a commentary on Sāmkhya in the classical Indian style of doing philosophy it would be perfectly in order to put this forward as the correct interpretation of the Sāmkhya aphorism which says that the manifestation of the hitherto unmanifested material principle (nature) takes place in the presence of

consciousness through a disturbance in the equilibrium of the three strands of energy. Anyway, we are told that given these two principles, both eternal and independent of one another, manifestation of the unmanifested starts.

Now the principle of materiality is constituted of forces that are very subtle and consequently not perceptible to the senses. It is through their combination in different proportions that manifestation of different orders of things takes place. As arising out of very subtle forces the order of manifestation also proceeds from the subtle to the gross in progressive stages, the actual material things being the last to evolve. The first category of manifestation is the intellect conceived as the principle of discrimination. This is the highest and the most subtle manifestation of the material principle and in it the strand of transparency, that which is conducive to physical manifestation, is predominant, the reason why it can function as the principle of discrimination. But however subtle the functioning of the intellect it is still a function of the material principle and as such unconscious in its functioning. Hence it will be a mistake to confuse it with consciousness which is aware of itself. This intellect is conceived under two aspects, the first is a cosmic aspect under which all discriminative activity found in nature is to be understood and because of this cosmic presence of the intellect nature shows unconscious purposiveness – like the flow of milk in the cow for the nourishment of the calf. Purposiveness then does not need an agent who consciously conceives a purpose, it is inherent in the subtle combination of material forces. Thus a Sāmkhya philosopher should not be surprised that computers function intelligently and purposively, only he would refuse to equate men with computers as computers are not consciously aware of their own activities. A computer can be programmed to write out 'I know what I am doing' any time it does something. This will still be a form of intelligent response, but it will not be the sort of self-manifestation that conscious experience consists of. However I see no reason why on Sāmkhya principles an artifact cannot be constructed which would be conscious precisely in the sense in which human beings are conscious. But of this more later.

This principle of intellect is also found in man and it forms the basis of all his mental activities. From the principle of intellect arises what is called ego sense, an unconscious principle of individuation which marks one thing out as distinct from another. In it the strand of drive and movement is predominant. It is this which acting in men makes

them feel their distinct individuality and stimulates and impels them towards various activities for the fulfilment of this individuality. It also brings the possibility of pain and suffering. Then further manifestation takes place. From the transparency aspect of the principle of individuation (all manifested aspects of nature have all the three strands of energy in different proportions) comes mind, the arranger of sensations into percepts and the source of volition, pleasure and pain. The mind is said to be an internal organ possessed of magnitude; it could well be the brain but the term 'brain' is not used in Sāmkhya terminology. (In Indian literature the mind is often placed between the eyebrows where the root of the brain is supposed to be. Sometimes it is located in the heart.) From the same source also arise five sensory organs of sight, touch, hearing, taste and odour as well as five motor organs like the hand, the feet, the speech organ, the excretory and generative organs. From the strand of heaviness or restrain inherent in the principle of individuation come five subtle and imperceptible principles of physical nature, principles which make possible the sensations of touch, hearing, sight, taste and odour. From these subtle principles arise various forms of gross matter, ether, air, fire, water and earth the very first entities (except ether) to be physically perceptible. Last of all from the combination of these in different proportions arise the common things of the world, rivers, trees, etc., and thus the manifestation of all possibilities inherent in the principle of materiality is completed.

There is a school of Indian philosophy called Nȳaya-Vaisesika which postulates atoms (these may be said to be infra-atomic, in some sense, since they do not, according to some authorities at any rate, have magnitude) as the constitutive elements of the physical and hence the material world. Sāmkhya believes that the term 'material' is a wider category than 'physical'. The physical universe evolves out of subtle material principles which make sensations of sight, hearing and so on possible and these principles may be compared to the 'atoms' of the other system. These lead to the formation of gross physical elements, air, water etc. and from the combination of these in different proportions arise solid substantial objects. But the 'material' includes much more than solid substantiality, it includes subtle although unconscious capacities for discrimination, individuation and so on. So the whole of the material realm cannot be derived out of physical principles responsible for the formation of perceptible and gross matter. Indeed gross matter itself evolves out of the principle of materiality and it arises at a later stage than discrimination and other mental functions. Inert, solid matter occupying space is therefore not the characteristic expression of the material. On the contrary it is a limited and comparatively late manifestation of the material principle which can function on a much more subtle plane. According to Sāmkhya, evolution of material characteristics takes place from the subtle to the gross rather than the other way. Sāmkhya would thus explain the development of the human organism in this fashion. All the subtle psychological capacities, capacities for intelligence, sense perception, action etc. are already inherent in the subtle material properties of the germ cell; the body as a physical object comes later as a gross abjunct of these capacities through cell division. Thus it is clear that everything that goes to make a man, except consciousness, arises purely out of the principle of materiality and insofar as this is so man is no different from other beings in nature. Human capacities of a characteristic kind, usually intellect, ego-sense and mind, are grouped together and referred to in the literature as the 'internal sense' as distinct from sight etc. which are external and this internal sense becomes the seat of the activities usually called 'mental'. These capacities constitute the internal sense as they function without any actual contact with an object, unlike the sense organs of sight etc. which are called external as they need contact. The external sense organs however are not the external physical instruments as such but capacities exercised through these instruments. The difference between internal and external senses, the mental and the physical organs, lies in the degree of the subtlety of their composition and the more the strand of transparency predominates in the composition of a thing the greater the subtlety. This is the counterpart in Sāmkhya of the modern doctrine of degrees of complexity and organisation in different manifestations of the physical principle of functioning.

Man with his internal sense, external sense and motor organs, and body composed of physical elements like earth, water and so on, would be complete as a functioning being even if we do not bring in consciousness which is altogether a different thing from all this. He can behave perfectly intelligently and purposefully making discriminations of all sorts and we can, by watching him perform, say that he is undergoing volitions, cognitions and feelings (Sāmkhya does not of course say this, but I find this implicit in its doctrines) but these would all be unconscious. That is, he would perform in a certain purposive manner without ever being consciously aware that he is performing in

such and such manner for such and such purposes. But the man as we know him now is conscious. How does he get to be conscious and what does his being conscious mean?

Now according to Sāmkhya there is no God and there was no creation. There are two eternal (i.e. uncreated) and absolute (i.e. neither derived from the other) principles at work in the universe, the principle of selfhood and the principle of materiality. The principle of materiality is sufficient to explain everything that there is in the whole realm of the manifested universe except for the fact of consciousness or knowledge that it exists. That is, it is capable of a great many things but incapable of self-manifestation insofar as it cannot become consciously aware that it exists. (The physical strand of transparency is also called 'manifestation' but this must not be confused with 'manifestation that' of knowledge.) And self-manifestation, or manifestation of itself as there, is the one function with which selfhood or consciousness is identified. According to Sāmkhya consciousness is inactive in every other way, and all activity springs from the principle of materiality. Indeed even self-manifestation is not to be thought of as an activity at all, it takes place through consciousness being what it is, and it is just the expression of the self's essence.

There are many selves according to Sāmkhya; this is understandable if we mean by 'self' the empirical man who is not just a consciousness function but has a body and a mind which individuate him and enable him to have empirical experience of which he can become conscious. But according to Sāmkhya 'self' means consciousness and it is only consciousness which constitutes man's real nature, not his physical and mental properties from which he can dissociate himself without any real loss. (Indeed dissociation leads to real gain not loss.) As self is nothing other than an identical function of self-manifestation it is not clear what makes for plurality of selves. Possibly these selves, even though identical in their essence as pure consciousness (consisting of nothing but a manifesting function) are existentially different, representing different possible centres of manifestation, and one centre of experience is not another, even if there is nothing in the nature of their experience, considered as pure consciousness, by which one can be distinguished from another. Anyway, selves are quite different from anything that belongs to the other pole of existence, that is, nature, including intelligence and other mental attributes, and there is no interaction between these two poles. That is, selves in their nature as pure consciousness are not involved in any of the activities of the manifested

universe and their function of manifestation simply consists in witnessing what goes on in the world including happenings in their own minds and bodies (or what are mistakenly thought their own). What we call an empirical self possessing a psychophysical organism as well as consciousness comes about through a special kind of relation called 'reflection' (as in a mirror) between the self and the intellect aspect of the organism.

The intellect, as we have seen, is constituted predominantly of the transparency strand of matter which is conducive to physical manifestation and although different from consciousness, its special nature as the most subtle matter makes it possible for consciousness to be reflected in it. In modern language the complexity of the structure of the brain as organised matter allows that certain of its activities exhibiting some special characteristics can receive in their medium another kind of function constituting awareness that, as a result of which conscious experience that such and such takes place. The intellect acts in conjunction with sense organs in receiving and discriminating sensations as a result of which the psychophysical organism undergoes unconscious cognition, volition and feeling. When a sense organ receives the impression of an object the intellect, according to Sāmkhya, gets modified into the form of the object and then as the result of the reflection of consciousness in the intellect the object is not just unconsciously cognised – this takes place at the level of the intellect – but also becomes manifested as that object. That is, one becomes consciously aware that an object and does not just cognise it unconsciously as would be the case if only the senses and the intellect were at work without consciousness function having been reflected in the intellect.

As there can be no interaction between the conscious and the unconscious Sāmkhya devises this special relation of reflection in order to explain conscious experience which an empirical self undergoes. But the trouble is that this reflection relation, unlike interaction, is hardly a relation by which the unity of an empirical self, who is not only conscious but has a body and a mind, can be effected and what happens is that Sāmkhya is forced to deny that there is such a unitary being at all. We have seen that both Husserl and Sartre deny that being a personal ego is a primary datum of our experience, which is just consciousness and not a person having consciousness. Nevertheless they also emphasise that consciousness must constitute a personal ego, and constituting does by no means imply postulating fictitious entities, it

only means that certain things are given at one remove which consciousness must grasp, by means of its intuitive and intentional functions. Sāmkhya says that the personal ego is not only not given in primary experience it is literally constituted by a mistake. The empirical man is not a unity of body, mind and consciousness as he thinks himself to be, he is consciousness reflected in a mind-body organism, but as his consciousness can be withdrawn or dissociated from the organism in which it is reflected it is a mistake to think that the organism constitutes an essential part of his being.

The mistake happens like this. Knowledge of objects (knowledge that) takes place, according to Sāmkhya, as consciousness is reflected in a certain psychophysical organism. As there is no interaction we cannot say that consciousness that gets reflected in a certain organism belongs to that organism or functions in conjunction with it. It just happens to be reflected there. But as reflection takes place consciousness, forgetting its nature which is purely that of manifestation, begins to think of itself as essentially consciousness of the object which has been manifested and as the discrimination of the object originally took place through the psychophysical organism self's identification of itself as the consciousness of an object also means its identification with the psychophysical organism. It is only through the psychophysical organism, internal sense and external sense organs, that objects are reached and discriminated and it is through the reflection of consciousness in this organism that conscious awareness of the object as such and such takes place. But consciousness appropriates as its own the whole of the medium and its functions in which it is reflected and it calls the whole thing 'I' rather than the function of consciousness which according to Sāmkhya is the real referent of the 'I'. In this identification it is helped by the principle of individuation or ego-sense which forms part of the internal sense of man and then it suffers the illusion that it itself is nothing but a function of the medium in which it is reflected, that is, the psychophysical organism. In its 'essential' nature however the self is just the knower or a witness who witnesses what goes on in the psychophysical organism but takes no part in its goings on.

Sāmkhya exhibits the full force of the tendency which Merleau-Ponty found in the Cartesian cogito. Having found that consciousness as a function of awareness is an independent and underivable reality, and naturally designating it as self, Sāmkhya goes on to identify the empirical self of man with his consciousness also called 'self'. When it says that it is a mistake for self to identify itself with the mind-body organism, this may be acceptable if by self we mean just the function of consciousness, but if we mean by self the empirical man then this assertion simply is not true. It may of course be said that the essence of man lies in his consciousness. This may mean two different things: (1) Man is nothing but consciousness, and this, empirically speaking at any rate, contradicts facts; or (2) man is distinguished from other creatures and things by the possession of self-consciousness. This second assertion is true, but it does not mean that man is nothing but his consciousness. Consciousness may be man's essence in the sense of being his distinguishing property but it certainly is not his only property. Again to say that consciousness is absolute may mean two things, (1) consciousness as a function of manifestation is not derived from the object it manifests and (2) the empirical subject X who is a conscious being is absolute and his consciousness of whatever he is conscious of is in no way dependent upon his psychophysical constitution or his environment; and this is manifestly false. The absoluteness of the principle of consciousness does not make man absolute unless we have identified him with this principle at the outset. If the principle of consciousness acts in man in conjunction with a mind-body system with which it interacts in some sense and through which interaction it gets modified, and if it is this whole interacting complex, consciousness-mind-body which is the empirical man, then the absoluteness of the principle of consciousness considered as a principle by itself would hardly make man absolute. But as Sāmkhya refuses to admit interaction while at the same time advocating dualism, it can do nothing else but say that man is identical with his consciousness and that his thinking of himself as an empirical ego which is limited by his mind-body organism is a mistake. This mistake can be corrected and man can, through right knowledge of his nature, dissociate himself from the functionings of the psychophysical organism and thereby attain an unconditioned status which rightfully belongs to him as self. Doing this is gaining freedom through liberating oneself from the limitations of the principle of materiality.

Thus the problem of freedom is here approached from a different perspective from Western philosophy. Nevertheless it is not altogether a different problem. We have seen in our discussion of the controversy between free will and determinism that exercise of free will, not in the sense of being able to do what one wants to do but being able to choose other than one did, depends on the ontological status of man as a

being who is not determined in respect of all his activities. It is this ontological status of man as a free being that Indian philosophy is interested in rather than free will as such. Sartre too is interested in this status of man, but he approached man as an empirical and contingent being who enjoys freedom only as such a being while in some Indian philosophy the tendency is to identify the whole being of man with this status and consequently to declare that all that is contingent does not form an essential or a real element in man's true nature. Anyway man is here conceived as altogether a different sort of being than is revealed in the determined functioning of the material world and his freedom lies precisely in realising this fact. It is clear that this different status of man is tied up with the fact that man has consciousness and since consciousness cannot be derived from unconsciousness it is argued that the *true* being of man is not conditioned by whatever belongs to the world of the non-conscious.

I think that Sāmkhya fails to give a satisfactory account of the concept of freedom *as it applies to the empirical man* and this is because of its identification of Self (the principle of consciousness) with self the conscious empirical man, both of which are referred to by the same Sanskrit term *Purusa* (Sanskrit makes no distinction between capital and small letters so this device can not be used to demarcate the two as in English). Again it is possible that Sāmkhya is misled by the epithet 'pure' which can be used to demarcate the principle of consciousness from an instance of the functioning of this principle as consciousness of this or that. The principle of consciousness is pure only in the sense that it is being conceived as a function in abstraction from empirical functioning which particularises it and gives it definite content. But the term 'pure' has also a moral connotation which suggests that whatever is pure is far more desirable than what is not pure. So consciousness as such, pure consciousness, becomes something with which men ought to identify themselves rather than with empirical consciousness which is relative to its content and hence not pure.

According to Sāmkhya the conscious empirical man comes about when this pure consciousness is reflected in a mind-body system via the intellect and when consciousness identifies itself with this system. With this identification pure consciousness has lost its sovereign independent status, for it is now consciousness as limited by the function of the intellect and the senses. The acceptance on the part of consciousness of its dependence on the intellect however is a mistake, for pure consciousness is by definition absolute. And so is man via his

consciousness. Man can regain his absolute status by ceasing to think of himself as a mind-body system. He is advised, in order that he may achieve this end, to adopt a deliberate policy of not treating what goes on in his psychophysical system as an essential part of his true nature. This will in the end bring about the required dissociation and consequently help man to gain freedom.

Now if identification of self with mind-body is false it is only false if we mean by self 'principle of consciousness' which is distinct from everything material by definition. It cannot be false for the empirical self as the empirical man would not at all be there if consciousness did not get reflected in a material system. It is therefore not true, as a matter of fact, that the empirical man is absolutely identical with consciousness. Sāmkhya of course thinks that the empirical man exists by a mistake. But mistake or not he is there – and if he were not there Sāmkhya would hardly have been concerned with his freedom – and the idea that he exists by a mistake can be explained by certain confusions one of which I have already mentioned.

But it is also rooted in the doctrine of a non-interacting dualism which leads Sāmkhya to formulate the relation between consciousness and mind-body in terms of something which is really a non-relation, that is, reflection. If this reflection relation were correct it would not have been difficult for a man to effect the dissociation that Sāmkhya recommends. For after all consciousness is conscious of itself and it remains aloof in its essential nature even while making a mistake, for if it did not, its nature would not really be transcendent; so it should be able to see that it is only being reflected in a system just as it is easy for a man to see that a reflection in a mirror of someone's face is not really his face, even if he does initially commit this mistake, when the distinct entity which is the face is presented in the original. Man, however, even when willing to acknowledge that his consciousness is not resolvable into his physical and mental attributes is unable to see that his physical and mental attributes are not genuine parts of himself as a being of a certain sort. In other words, he finds that what he calls his consciousness functions in conjunction with his mind-body system and is not totally divorced from the affairs of the latter so that the relationship is in fact a closer one than that of mere reflection. If his consciousness manifests not everything but only that which relates to his psychophysical organism then its manifesting activity is in some sense dependent on the system whether or not the very function of consciousness is derivable therefrom.

Of course if consciousness as a function is underivable from the psychophysical organism then man is mistaken if he believes that everything about him can be explained in terms of physics and chemistry, but not if he believes that the system in association with which his consciousness functions is genuinely a part of himself as much as consciousness. It may of course be said that the empirical man is identical with self in essence. But the term 'essence' only refers to what is distinctive of man, it does not refer to his total being. And if one says that that which constitutes the essence of man is the real man one is only making a value judgment in the guise of a factual assertion, for what one is saying is that mind-body does not count or is not important. It is not as if in man consciousness and mind-body go their respective ways and there is no unity. Even according to Sāmkhya, although consciousness and mind-body do not interact there is a sort of pre-established harmony between them as a result of which one can be reflected in the other. Consciousness is not reflected in a system which has not got the subtlety of composition of the psychophysical organism that man is, and if we use this reflection language it is this total reality, consciousness reflected in a mind-body system, which makes what we call the 'empirical man'.

A conscious organism is limited and conditioned by laws which regulate its functionings. But Sāmkhya wants an absolutely unconditioned status – a state of complete freedom from all limitations – for man which, it believes, man can have because of his nature as transcendence. But it is not limited and piecemeal transcendence of given conditions as in Sartre, it is total transcendence of all given conditions effected through disowning the organism through which conditions affect man. Hence the identification of Self with self. If this disowning can be done the empirical man will no doubt become a Self but there will be an end of man as we know him. This for Sāmkhya is no great loss since it sees human life as full of frustrations and sufferings. Indeed it is because Sāmkhya wants freedom from frustration and suffering that it proclaims the total freedom of man which is paradoxically both a fact here and now and an ideal to be achieved. Freedom for man then means in Sāmkhya transcendence of all limiting conditions as well of all sufferings and frustrations that go with these by a total denial that these things have anything to do with man's real nature. This man can do because he is a conscious being and consciousness transcends the world (or in Sāmkhya language is totally dissociated from the world). The Sāmkhya ideal is total transcendence, of being

one with pure consciousness which is supposed to be tre true nature of man and it is only by becoming this that man can be truly independent and free. This of course, as I have remarked earlier, means, in effect, ceasing to be man. But what if a man does not want this freedom, what if he wants to live within the limitations that go with being a part of the material world? To tell him, as Sāmkhya would, that he wants this only because he is mistaken about his *real* status as an unconditioned being is only to reveal Sāmkhya's own presuppositions as to what is worthwhile. But what is not worthwhile to a Sāmkhya philosopher may be worthwhile to someone else who does not particularly care about being totally unconditioned.

But Sāmkhya may have a point shorn of the exaggerations to which all systems are prone, when it insists that man's essence lies in realising a state of ontological freedom which consists in a conscious experience of oneself that one is not, or at least not altogether, (Sāmkhya says not at all) a derivative of material forces and that one does transcend given conditions in some fashion. Only, this truth is a partial truth about the empirical man and not the whole truth. As consciousness is or can be a function of transcendence a man who believes himself to be free (in this limited sense which requires that man has to assert his freedom in order that he may be free) can, even whilst he is being limited by various forces including his own impulses, refuse to believe that his being is dissolved into these limitations. If a man does this he realises his distinctness as a conscious being who can, even if only at moments, assert his independence and gain it in some sense simply by asserting it no matter what are the limitations by which his life is hemmed in. As Sartre says, the for-itself is not glued to itself, but unlike Sāmkhya Sartre does not stress the intrinsic value of a state of realisation that one is a function of transcendence, and freedom is important for him primarily in the context of realising various empirical possibilities. According to Sāmkhya a feeling of freedom is an end in itself since it is realisation of man's essential nature. Sāmkhya of course fails to make the distinction that needs to be made between the concept of essential nature which symbolises only what is distinctive of man and his total being which includes things common between him and the non-human world and so it defines man as a totally unconditioned being. But if one can, without agreeing with Sāmkhya as regards this definition, stand aside and witness happenings even while they are happening to oneself, refusing to believe that one's being can be measured in their terms, one can realise in oneself, in relative terms, a

feeling of freedom from dependence on things external to oneself that Sāmkhya thinks is a desirable state of mind for a man to be in. It is desirable because it realises the ontological possibility, on the part of man, of experiencing a dimension of his own being which tends to get lost in the pressure of limiting conditions. This dissociation of course is neither complete nor can it be achieved once for all as man happens to be both conditioned and unconditioned. Anyone who finds this realising of oneself as unconditioned a desirable state can carry dissociation to its possible limits, but it depends on what one wants.

The human mind in all its contents is conditioned in every specific case by the particular balance of material forces within the character or disposition of the given individual – his perception, thoughts, feelings are but the functions of his particular manner of being as a psychophysical system. When consciousness is thought to be just an attribute of the mental a man remains determined by his conditioning. When consciousness stands aside and exercises this function of dissociation or transcendence it becomes possible for man to realise that he is more than a conditioned being, since he can question what is going on or disown it in thought. The Sāmkhya doctrine of total freedom does not apply to the empirical man and it comes to be advocated through a confusion. What I am saying, and here I am reinterpreting Sāmkhya, is that man can have moments of realisation when he is able to enjoy with respect to himself a feeling of freedom which comes from not considering oneself a function of one's conditions, psychological, physical and environmental, despite the fact that what one can accomplish remains limited by these conditions. This possibility of enjoying apparently contradictory feelings about oneself comes from man's being the meeting ground of two contrary principles both of which express his real nature. One is of course free to accept one or the other component of one's being as being more 'real' (more worthwhile, desirable, pleasurable etc.). But this is a value judgment rather than a descriptive one and the making of this value judgment depends on one's taste and temperament.

Man's feeling himself free does not imply that he can do what he likes, for his doings are not free, or are free only in a limited way. Being free is a state of experiencing oneself as a kind of being the being of which is not determined like the being of things. One may of course say that this freedom is empty if it accomplishes nothing and it makes no difference whether man has it or not. Now, it may be nothing in

terms of doings (although this is not wholly true as I shall discuss later) but it is not nothing in terms of being, of being a kind of being who experiences oneself, as not being just a set of fixed reactions and who transcends, in some fashion, given conditions. From the point of view of being, of self-experience, man's accepting himself as free *is* his being free and there is no sense in saying that a man may be mistaken in this, for here no claim is being made as to what one can accomplish, it is only about what one is aware in oneself of being. And the value of being unconditioned in this manner lies precisely in enabling man to feel unconditioned.

For an empirically minded philosopher all this is no doubt meaningless talk, for it makes no difference either way whether one is free or not. This criticism would be justified if man were to be conceived as a bundle of doings, how one thinks or feels about oneself being totally irrelevant in the scheme of human living. This however takes no account of the fact that man is a self-conscious being who can be aware of himself as being such and such nor is it borne out by evidence that it makes no difference to man whether he thinks of himself as free or not. Many people resist the suggestion that they are in no way different from material things (except of course being much more complex) even when this is not supposed to affect in any way their capabilities and achievements which may even be enhanced by a better grasp of mechanisms involved in physical functioning. This is because, to them, what they think of themselves is important, and as Sartre says, being self-conscious they can hardly fail to have a certain comprehension of their capacity for transcendence. People resist this suggestion because they are aware of themselves, however vaguely, as having a different status from material objects and it is out of this awareness that they protest against the doctrine of scientific determinism which assimilates them to objects.

It may be said in criticism that if feeling unconditioned makes one unconditioned then feeling, say, honest should make one (who is in actual fact dishonest) honest. But there is a vital difference between the two cases. Being honest is tied up with acting in specific ways, and if someone who feels himself honest does not act in these ways it may justly be said that he is wrong in his feelings. But feeling unconditioned by virtue of the fact that one believes oneself to transcend given conditions in some fashion, is not a claim about being able to do such and such things (although feeling unconditioned may actually lead to greater achievement). When this feeling just expresses a way in which

one experiences one's manner of being in the world without any reference to one's performances, it can hardly be something about which one can be mistaken as one can be about one's own honesty. It may be said that man may still be nothing but a highly organised material system for all one's feelings to the contrary. Well, it seems to me that the very fact that one feels oneself to be not just this shows that one is not just this. If there is a sense in which this kind of feeling may be said to be a function of the working of material laws then it is unimportant whether man is nothing but a material system or not for it detracts from nothing that he now thinks makes him different from matter including his dissociation and transcendence functions. For matter has here come to possess properties which are ordinarily believed to be, in a very broad sense, spiritual.

Now, Sāmkhya analysis of the relation between consciousness and mind-body as being that of mere reflection presents other difficulties. Sāmkhya dualism is most radical and consciousness in this system remains a mere observer of what is going on; it never becomes an actor. It is neither modified by the system in which it is reflected nor does it modify the functioning of the system itself. In Sāmkhya language it remains a mere witness, isolated and independent. The mind-body system functions entirely on materialistic principles involved in the three strands of forces, their combination in different proportions and the regulation of the forces of lightness and movement by the third called heaviness into law-governed behaviour. Nowhere in the functioning of this system is the agency of consciousness required to intervene. Indeed consciousness must not be thought of as an agency of any sort. There is not even any change going on in consciousness, for all that consciousness does is to manifest what is going on in the world, both mental and physical, and it does this by being what it is, a manifesting function.

This makes it impossible to understand how what is called a free action – on the possibility of which depends the occurrence of moral and spiritual behaviour – is possible. A moral act may require one to check some of one's natural tendencies which if not checked may result in some undesirable state of affairs. How can this kind of action come about? Take, for instance, the Sāmkhya advice that man should deliberately dissociate his consciousness from the states of his mind and body when there is no natural tendency in man to do so. What is it in man that can effect this? The intellect or mind, being unconscious by nature, can hardly bring this change about and if this is to be brought

about at all it can only be brought about by conscious effort, and this means that consciousness does not remain a mere witness but actually does something about what goes on. It may be said that consciousness brings this dissociation about not by doing anything in the way of influencing the functionings of the mind-body system. All that is necessary is that consciousness should be aware of the need for dissociation and it will come about merely through this knowledge. Unfortunately the matter is not so simple and, as Sāmkhya is well aware, it needs tremendous effort and prolonged discipline to gain what Sāmkhya calls 'freedom'. Such effort is possible only on the part of a conscious being who actively undertakes to refashion his natural tendencies and this requires freedom of will.

As a matter of fact determinism in the form of law of Karma (action) is a universal article of faith in Indian thinking and this determinism is even more pervasive than in Western philosophy inasmuch as even man's heredity and environment are here determined and are not mere accidents. The law of Karma is a causal principle which extends over more than one life. A man is what he is in his present life, his mind, body and environment are as they are, because he has acted in a fashion in a previous life, which through the mechanism of habit formation has created in him tendencies of a characteristic kind and these in their turn have determined the kind of mind, body and environment he should have in order that these tendencies may be worked out in this life. This means that he acts as he does because he is determined to do so through having a certain kind of mind, body and environment. This appears to leave no scope for freedom of will, as how a man shall act depends on certain given conditions. This determinism however is not the whole truth about man and it does not or need not regulate the whole of his life. As K. Bhattacharya[1] points out, the law of Karma determines three things in man (1) his physical and mental make-up (2) span of life, and (3) the amount of pleasure and pain that he will experience and the direct and indirect causes of such experiences amongst which are included environmental factors as well as his actions, but not all actions. Actions are of three kinds. (1) Purely biological and reflex actions; these are determined and can hardly be called actions, (2) actions induced by emotions of attachment and repulsion (there are many varieties of these). An individual feels these actions to be his actions but they are not free since they arise

[1] The Status of the Individual' in *The Indian Mind*, ed. by K. W. Morgan (Ronald Press, N.Y., 1933).

out of certain given tendencies. These are the two varieties of action that are determined by the law of Karma. But there remains a third variety of action possible for a man to perform. This is moral or spiritual action and it takes place when an individual has ceased to be guided by his innate tendencies of attraction and repulsion and has chosen instead to act out of considerations for what is desirable. This sort of action is called a detached action. It is not determined by pre-existing psychological tendencies, on the contrary it is only when an individual has learnt detachment from such tendencies that it becomes possible. Practice of such tendencies gradually builds up a disposition of detachment in man. This disposition does not bind man to his impulses as do tendencies of attachment and repulsion. Man is thus determined, but only so long as he has not learnt to act in a detached manner. With detachment comes freedom.

How are free actions possible in the Sāmkhya scheme of non-interacting dualism? Sāmkhya requires that we undertake detached actions in order that we may dissociate ourselves from our innate tendencies with which we normally identify ourselves. If all our actions were based on innate tendencies detached actions would be impossible. They are possible only when consciousness intervenes and controls tendencies rooted in natural attraction and repulsion. This control cannot come about through the functioning of the mind-body system which according to Sāmkhya is unconscious and acts according to laws of material functioning. But it is urgent for Sāmkhya that free actions should take place, for man cannot be in a position to realise his unconditioned status unless he has learnt to act freely. It is no good saying that man is eternally free and he has only to realise this to be free as Sāmkhya commentators are inclined to say. This realisation needs effort and there must be an agency in man which can initiate effort. I do not see what this agency can be if not consciousness.

Sāmkhya makes consciousness wholly inactive and a pure witness in order to ensure its autonomy. This autonomy however does not require that consciousness must be wholly ineffectual in bringing about changes in one's mental dispositions. Of course there is the traditional difficulty of explaining how two wholly independent principles which have nothing in common between them can interact and this difficulty is not overcome by postulating a relation called 'reflection'. As consciousness is radically different from the intellect it is equally difficult to understand how consciousness gets reflected in the intellect. Intellect no doubt is conceived under the model of physical light

which is conducive to manifestation. But physical manifestation is something quite different from the 'manifestation that' of consciousness. As Sāmkhya admits, there is a pre-established harmony, a relation of fittingness, between consciousness and the intellect which allows for reflection. Why cannot the same fittingness allow for influence? When consciousness influences the mind-body system it does not remain a pure witness but it need not lose its capacity for this witnessing function simply because it does not always remain a witness. So autonomy need not be lost despite interaction.

It may still be felt that if consciousness and mind-body represent two radically distinct principles interaction between them is not possible. I do not know if this must be so. But if this is so, there remain three alternatives. (1) To say like the materialist that consciousness is not a different principle at all but is just a function of matter. This however fails to explain, for some people at any rate, some of the capacities that we associate with being human. (2) One could say that the material principle is not really a distinct principle but is only a modification of consciousness. For most people this appears even more implausible. (3) We could say that the consciousness principle and the material principle belong at different points in the same scale of some fundamental energy system of which these two are manifestations at different levels of existence. As levels of existence are qualitatively different manifestations are different too and this means that as a distinct level of existence one cannot be reduced to the other. As we have conceived of some fundamental system of energy which contains the possibility for consciousness as well as material functioning interaction would not be impossible although consciousness would not be a derivative of matter nor matter of consciousness in this scheme, both being equally valid expressions of what there is. It is of course difficult to see what this fundamental energy system can be like and we shall have to be satisfied just by describing it as 'that which makes both conscious and material functioning possible'. If intelligent discrimination is a function of the material principle – although much more subtle than non-intelligent functioning – as Sāmkhya says, and as recent researches indicate, then there is no reason why, as we go higher up the scale of subtlety of energy manifestation, we cannot come upon conscious manifestation of that discrimination as so and so. But so long as we define consciousness as that which manifests itself and unconsciousness as that which cannot do so there remains a radical distinction between the two concepts.

Sāmkhya subscribes to a doctrine of immanent causality which holds that the possibility of the whole of the manifested world is contained in the three strands of energy and their different possible combinations. Not only does at no stage something arise out of nothing, it arises out of its own potentiality already there. Different manifestations take place only when conditions become suitable for the manifestation of energy in these forms. If so, the late arrival of consciousness on to the scene of the world does not mean that it has been produced out of unconsciousness. It only means that conditions for the manifestation of consciousness in this form (as found in man) is now ripe, but the possibility is already contained in the primal source of energy which has now been conceived to include both consciousness and matter. If so, the functioning of consciousness in conjunction with the mind-body system is not a mistake but a distinct possibility.

The Sāmkhya doctrine of pure reflection may be modified by saying that consciousness interacts with a given mind-body system only when the mind-body system is subtle or complex enough to reflect consciousness in it. If this is acceptable, then the question which Hilary Putnam asks in his article 'Robots, Machines or artificially created life?[1] – whether robots can be conscious – becomes to my mind answerable. Robots are artificial machines which can simulate human behaviour to such a degree that one could by watching their activities say that they have 'thoughts', 'feelings', 'attitudes' etc. in the same way as we can, by watching people behave in such and such ways, say that they have thoughts. Do these robots have consciousness? Now if consciousness simply means the ability to behave in such and such intelligent fashion then of course whatever shows such ability is conscious. But are they conscious in the human sense, that is, conscious of being conscious beings? There is no evidence that robots are conscious in this sense although there is evidence that they are conscious in the sense of being intelligent. It is difficult to say what this other evidence consists of but intelligent performance is obviously not enough. At the moment we believe that human beings have consciousness, where consciousness includes a capacity for self-manifestation, whereas robots do not – this is why they are called robots rather than human. But refusal to call them human is not just a matter of prejudice against the hardware of which robots are made instead of flesh and blood, as Hilary Putnam suggests. It happens to be the case that we do

[1] *Philosophy of Mind*, ed. Stuart Hampshire (Harper & Row, New York, 1966).

not believe that we have evidence that robots are self-conscious in the way human beings are. As soon as we have unambiguous evidence that robots are like humans in this respect I do not think that there would be any genuine difficulty in classifying them with humans on the score that they are made of hardware. We would perhaps think of them as a distinct racial group – it is no longer just a matter of colour and features as with present racial distinctions but of body material as well – and we might even be prejudiced against them because they possess this different body material but I do not think that we shall really doubt that they are human in the appropriate sense if they do give us the required evidence. There are some very primitive tribes who are very different from advanced white races of human beings not only in colour and features but in their capacities for performance as well and who would possibly fall considerably below robots belonging to an advanced technological society in this last respect, but nobody seems to entertain any doubt about their being human, although there is no lack of prejudice. So I do not think that human unwillingness to accept robots as human in the relevant sense is all a matter of prejudice against hardware as Putnam suggests.

If robots begin to give evidence of being aware of themselves as conscious beings then according to Sāmkhya principles the mind-body system that a robot is is subtle or complex enough for consciousness to be reflected in it in the same way as the neurophysiological organism of human beings is. The fact that it is made of hardware rather than of cells then obviously makes no difference. But it may well be that the physical and chemical properties of the neurophysiological organism, which enable consciousness to be reflected in it, are not present in hardware. But this is an empirical consideration which keeps the possibility open, in principle, of artificially creating machines or organisms which will be conscious in the human sense.

Some philosophers have objected to the proposal that we should accept robots as conscious on grounds which Putnam thinks are inadmissible. The fact that robots are machines makes people feel that their functioning is as simple as the pressing of a button and the appearance of a light. As a matter of fact, they have a complicated internal organisation in which one state is linked with other states just as human internal states and processes are. This cuts out the possibility suggested by some that robots, unlike humans, can be reprogrammed in any way we like. This also makes, according to Putnam, the robot utterance 'It looks as if there is something red in front of

me', when his internal state Flip Flop 72 is on, not unlike a human utterance of the same nature when his optic nerve is stimulated in a certain fashion. Kurt Baier however objects that these two utterances should be taken as evidences of same abilities in robots and humans. In the case of a robot the sensation evokes the utterance, in the case of a human being the connection is not merely causal and the utterance is not evoked by the sensation. A man utters these words because he knows that he is having this sensation. Putnam replies that a robot is just the same – the connection is not merely causal and the sensation does not just evoke the utterance. The robot has to be in an internal state which realises the predicate 'knows that *p*' in order that it may utter a 'psychological' statement 'It looks etc.' Baier however thinks that there is a difference between this robot utterance and my utterance which I utter because I know that I am having a sensation which implies that the robot does not know this.

Putnam makes short work of this objection by interpreting 'knowing that' purely as being able to undergo a process of sensation or behave in a certain way. If knowing that one is having a sensation is identical with having a sensation then the robot too knows that. If knowing that consists in being able to describe then the robot knows that too. The psychological-physical distinction is applicable to robots if it is applicable to humans. A robot can say 'It looks etc.' when it does not know that Flip Flop 72 is on or even the existence of Flip Flops. The robot can even say 'I know that I am having a sensation'. These utterances are instances of differences between mental and physical states and this distinction applies to the functioning of robots in the same way that it applies to that of humans.

I agree with Putnam that both a human being and a robot will have to be in an internal state which realises the predicate 'knows that *p*' in order to utter a psychological sentence and that the distinction between mental and physical is applicable to both. It also seems to me that a human being can utter a psychological sentence of the form 'It looks etc.' in the same way as a robot without having to know that *p* in any sense that is special to him (that is, involving more than an internal state that realises the predicate 'knows that *p*'). It is not necessarily the case that I utter this *because* I know that *p*, if my knowing that is something over and above my having a sensation, and my utterance may be an exact parallel of that of the robot. But it equally seems to me that it is possible that on some occasions I utter this because I know that I am having a sensation in a sense in which a robot,

as currently conceived, does not. Here I am aware of being conscious that I am having a sensation, and I suspect this is what Baier has in mind when he objects that the robot does not know that *p*. We do not believe that robots experience this sort of thing although there is a perfectly good sense – as expounded by Putnam – in which a robot may be said to know that *p*.

Putnam may say that I am making this distinction because I am prejudiced against robots. But hardware does not worry me and for my part I am perfectly willing to concede that robots can have self-consciousness in the full human sense as soon as I come to believe that I have evidence that they have it. But I need for this purpose more than intelligent performance of the type robots currently exhibit. I manage to have an intuitive understanding of human beings as humans who have private experience and I do not have to judge this by watching their performances all the time. Nor do I deduce this understanding about human beings from their appearance and body material although I associate a certain appearance and a certain body material with being human. There is no reason why this sort of understanding should not operate just as well in the case of robots. Of course, as I said, if a human being is conceived as consciousness reflected in a highly organised material system with which it interacts then there is no reason why it should not in principle be possible to artificially create a conscious being by creating a material system of the requisite degree of complexity of matter and form. As to whether we shall ever manage to accomplish this, that is another matter. But I just do not think that robots as currently conceived have self-consciousness, so they do not yet qualify for being human.

VEDĀNTA

The monistic Vedānta system of Samkara, henceforward to be referred to just as Vedānta, says that in the ultimate analysis there is no such thing as an independent material principle for consciousness to interact with. It is we who through ignorance impose the idea of an inert material world and its manifold distinctions on to the true nature of things which is pure consciousness. This material world of course exists as long as this state of ignorance exists and at this level some sort

something or other, a distortion of the true state of affairs and so forth, all of which seem to mean that the fact concerned is unimportant and not worth bothering about. This applies to philosophy of mind just as much as to any other problem and we are often offered, at least in theory, a vision of man, who is either just brain and nervous system or nothing but pure consciousness, these two being difficult to put together in a unitary scheme.

Man however may be a complex being, and it is possible that a full understanding of how man does or can function requires that we have to approach the question from different levels or points of view, none of which can claim to possess the sole truth about him. This may be intellectually untidy but nearer to facts about man as we experience them. I am not however saying that philosophers ought not to have preferences or that preferences felt by them cannot be in any way justified. But this justification seems to me to lie in showing that one can hope to achieve some worthwhile result by acting or thinking in the line of these preferences rather than in simple statements of fact as to what is the case, when such statements obviously contradict aspects of our experience of ourselves and others. Behaviourist-materialist preferences can be justified in terms of our scientific understanding of human behaviour insofar as this is determined by brain states; phenomenological-existentialist preferences in terms of possibilities of man as a responsible being who can acknowledge his responsibility to create his own mode of existence insofar as this can be done by a self-conscious understanding on the part of man of his own performances; and Indian preferences in terms of achieving a transcendental dimension not disturbed by the ordinary limitations of the human condition insofar as this can be realised by the witnessing function of consciousness. All these objectives may be worthwhile, but none of them can claim to be exploring the only true or real nature of man, if man, as I believe to be the case, is a multidimensional being and it lies in his nature to be capable of functioning in all of these modes.

I

The Materialist-Behaviourist View

CONSCIOUSNESS

The materialist theory of mind – with which behaviourist theories are associated – may also be said to be a theory of no mind in some sense. Most people, when they talk about their minds, as distinct from their bodies, vaguely consider mind to be a distinctive kind of functioning which is not thought to be identical with the functioning of the body, although they certainly also believe that the functioning of the mind is intimately associated with the functioning of the brain. Even theories of mind which insist on the distinctness of mental functioning associate the occurrence of mental states with the occurrence of brain states whether the relationship between these two kinds of states is conceived under the model of parallelism or interactionism. The materialist theory denies that there are two distinct types of states at all to interact or to be parallel with one another – one occurring on the internal screen of the mind, as it were, the other in the publicly inspectable brain. There is just one kind of states, the brain states, with which the so-called mental states are identical.

Of course the materialist account is not that clear-cut. Sometimes materialist talk about mind is such that we could say that there are mental states but somehow these states are identical with brain states; at other times it looks as if the materialist is denying that there are any states occurring on their own in a psychical sphere which are at all identifiable as mental states. D. M. Armstrong in his *A Materialist Theory of the Mind* appears to hold this latter view and as this gives a plausible account of what are normally called 'mental states' I shall base my analysis mostly on this variety. But of course even Armstrong has to admit that there occurs consciousness, or that there is such a state as being conscious, which does not fit in very well with the

doctrine that there are no mental states. But if we overlook consciousness it seems to me that Armstrong's doctrine that mental states are just brain states has a great deal of plausibility. To say that mental states are brain states is not to deny however that the logic of the language of mind is different from the logic of the language of brain, it is only to say that these two different languages do not have two different kinds of states as referents. Talk about mental states is justified only insofar as the physical states themselves are approached under a different set of concepts with a logic of their own. So the terms 'mental state' and 'brain state' refer to the same state but with two different purposes in view.

I shall no doubt be criticising the materialist view of mind, but there is a good deal in it with which I agree, so it would perhaps be wise to state at the beginning where I agree and where I do not. As far as I can see, the concept of mind refers to two distinct types of occurrence: one is a state of being consciously aware of something, and this is privately experienced, the other is a characteristic type of activity, such as moving a class piece when found necessary, which is an intelligent and thoughtful response adequate to what is demanded by the situation and which shows a dispositional readiness to make certain distinctions as and when the situation demands. These and similar sorts of activities including use of speech are called 'mental', but they are not necessarily accompanied by any privately undergone experience of conscious awareness that such and such is the case. So here the concept of mind denotes intelligent, discriminative capacities of a performatory and dispositional sort which may function perfectly adequately without giving rise to a conscious experience that such and such is the case. Animals have a mind in this sense – although their mental capacities are much below those of humans – without having a mind in the first sense which signifies the presence of a consciously felt private experience.

These two uses of 'mind' when applicable to humans are so interwoven that it is difficult to disentangle them. Speech, for instance, may be used without any private conscious experience that one is doing so and we do not need to have any privately experienced conscious awareness of what we are talking about in order that we may use speech intelligently as a performatory response (when engaged in conversation, for instance). Nevertheless the use of speech is often intermingled with conscious awareness of what we talk about so that it is difficult to separate performance from conscious experience in a thoughtful

use of words, and consequently performance seems to be only an aspect of conscious awareness or vice versa. But it is clear from the examples of animals that to have a mind in one sense – that is, mental_1 – is not necessarily to have a mind in the other sense as well – that is, mental_2 – where it denotes a capacity for conscious experience. But a great deal of confusion does arise owing to the fact that language which is perfectly adequate to mental_1 – a sense in which both animals and humans have mental capacities – is used in respect of mental_2, and since both sorts of capacities are called 'mental', which are supposed to characterise what is called 'the mind', it is not properly acknowledged that mental_2 has characteristics absent in the case of mental_1. The no mind theory, for instance, is perfectly understandable if it means that nothing non-physical need be involved in the capacity of animals and humans to perform discriminative, intelligent activities of various degrees of complexity, including use of speech in the case of humans, but it becomes difficult to understand when a state of conscious awareness is characterised as nothing but a physical state.

I accept the no mind theory (meaning that there is nothing non-physical at work) in respect of intelligent performances and dispositions and if I use the term 'mind' to denote something non-physical, as I would no doubt be doing, it would mean the sphere of conscious awareness. So what the materialist, such as Armstrong, says about mental states, that they are just physical states, is according to me perfectly acceptable if we make a distinction between the mental and the conscious – that is, mental_1 and mental_2 – the term 'conscious' being applicable only to an experience of conscious awareness that *p*, which is a felt state of awareness and which is identifiable as such by the person having the experience. The logic of the term 'mental' (meaning mental_1) is such that it is applicable to any functioning of human beings which involves intelligent discrimination between stimuli of different order, which is intentional (aimed at or directed towards something and not random), purposive (goal guided) etc. but which may not involve a state of conscious experiencing. It is easy to see that when 'mental' is defined in this way, there is no state which is mental, 'mental' signifying an order of happenings which is existentially different from the physical. The use of the term 'mental' here only denotes certain dispositional characteristics of brain states, such as intelligence and purposiveness.

I want to emphasise however that to talk about mental_1, is not necessarily to talk about mental_2 and mental_2, or the conscious,

meaning a state of felt awareness that such and such is the case, is not identical with the physical, at least not in the straightforward way that performance may be said to be. This is because felt experience has an existential identity of its own precisely because it is felt, which mental_1 in the sense of performance or disposition has not, and this difference remains there even if a state of conscious experience could be produced by a physico-chemical state of the brain. The materialists however, having shown that mental_1 is identical with the physical, go on to include mental_2 under mental_1 without any consideration of the fact that mental_2, conscious experience, has a felt existential identity of its own which mental_1 has not. It is this part of the materialist doctrine, that a state of being conscious is identical with a brain state, as well as the doctrine which some put forward that there is no such thing as experience, with which I disagree.

It is not customary in Western philosophy to make any distinction between 'mental' and 'conscious' and consciousness is generally thought to be an attribute of the mental. This was not confusing as regards the distinctiveness of the conscious so long as the mental was not identified with the physical, although of course it meant that all mental activities, intelligent performance etc. were thought (wrongly I believe) to have mental_2 characteristics without which, it was thought, they could not occur. Now that mental_1 is increasingly being shown to be a characteristic of the physical, that is to say, an outcome of physical and chemical properties of brain states, it is less confusing to make this distinction explicit. Anyway, I suspect that the opponents of the materialist doctrine and consequently of the identity thesis object to this identification not because they think that intelligent performance and dispositional characteristics involved in such performance cannot be identified with a brain state but because they think that experience or conscious awareness that *p*, usually referred to as 'raw feel of experience', cannot be. The materialists often respond to this by showing that much of human behaviour in respect of which it is customary to apply the epithet 'mental' just does not involve any raw feel. This seems to me perfectly true. But it also seems to me that raw feel of experience does occur at times, even if we can explain performances that may go with this experience in physical terms, and a piece of behaviour which involves raw feel of experience cannot be assimilated without a residue with behaviour that does not. Materialists who admit to conscious experience, however, claim that conscious experience, like intelligent performance, is a brain state so there is after

all nothing but brain states involved. This is where people who insist on the separate identity of mind (defined in terms of consciousness rather than performance) differ, for it is difficult to see how a thing like conscious experience can be produced by electrical discharge in nerve cells of the brain, and even if it can be, this still does not get rid of the existential identity of conscious experience felt as something to which the term 'physical' does not seem to apply.

Armstrong however assimilates a state of consciousness with a mental_1 state such as is involved in our ability to perceive, say, a colour.

> If I perceive a physical situation, then we have an inner mental state 'directed' in a certain way towards a certain physical situation. . . . Now, if I am aware, not only of the physical situation, but also of the fact that I am perceiving, then we have a further mental state 'directed' in the same sort of way . . . towards the original mental state. And if this further mental state, which *qua* mental state is simply a state of the person apt for the production of certain behaviour, can be contingently identified with a state of the brain it will be a process in which one part of the brain scans another part of the brain. In perception the brain scans the environment. In awareness of the perception another process in the brain scans that scanning.[1]

So for Armstrong the difference between a mental_1 state and a state of being consciously aware can very well be explained in terms of the different functions of the brain states involved in these two different cases. As he says, in case of perception of the environment the brain state scans the environment, in case of awareness of perception the brain state scans the inner state. But this is merely an interpretation that has been put on the functionings of the brain with reference to the ends they are being said to serve. The scanning process is a functioning of the brain and as such a physico-chemical process. As such a process it is the same whether it is supposed to scan the environment or an inner state. If the physico-chemical process involved in the scanning of the environment does not generate conscious awareness then it is difficult to see how the same type of process involved in the so-called scanning of an inner state does this. If the difference that appears between conscious awareness that perception *p* and perception of *p* without conscious awareness, is to be explained by two different scanning activities of the

[1] D. M. Armstrong, *A Materialist Theory of the Mind* (Routledge, London, 1968), 94.

brain then these two scanning activities must be shown to be different in appropriate ways. Simply to say that they do different jobs, one scans the environment the other an inner state, will not do, as this just states the difference as one intends it in terms of functions without explaining how this difference comes about. The difference then must be explained either by the difference in physico-chemical functioning of the two states, or by their different localisation. But neurologists do not believe that the physical process involved in conscious awareness is localised in any particular area such as seeing or hearing is. Nor do they believe that the physico-chemical process itself undergoes change when performing different functions. The processes in the nervous system, whether they are accompanied by consciousness or unaccompanied, show no difference in the chemistry of the nerve cell and the physics of nerve conduction, and no unique physical or chemical process has been found which is associated with the generation of consciousness. According to Sir Cyril Burt the brain is indeed a highly complex mechanism, but it suggests a mechanism designed to detect or transmit conscious activity rather than to generate it. The logic of neurophysiological functioning then does not force this contention upon us that consciousness is a brain state.

Anyway, it is only with regard to conscious experience that I differ from the materialist account of human behaviour, not with regard to its contention that the mental is the physical when mental means $mental_1$. There is a dualism, according to me, in the functioning of a human being, but it is better described as dualism of conscious and mental-physical than as dualism of mental and physical. For when the 'mental' means $mental_1$ there is no dualism. Whenever I am criticising the materialist account for saying that the mental is the physical I am doing so only insofar as 'mental' is made inclusive of $mental_2$ and is not kept restricted to $mental_1$.

Non-materialists such as the interactionists seem to me to give consciousness a greater causal role than it does have in human performance for it is likely that most human performances, including thinking *as manifested in intelligent use of words,* can be explained purely by brain function without the aid of consciousness. But it does seem to me that consciousness has a role in human performance in the case of actions which involve conscious choice and decision between alternatives, and where action that follows follows because it has been chosen by a conscious process. Of course even here consciousness does not produce the action, it is produced by brain process; but it has a role in influencing

the initiation of the particular brain process involved in such action, for if consciousness did not play a part in the form of choice some other brain process could conceivably be there resulting in some other action. How such influencing takes place is a mystery insofar as we do not know of such influencing of the material by the non-material anywhere else in the functioning of the universe, and the account of human behaviour can be made much simpler by saying that it is only brain processes which influence other brain processes, consciousness being simply an effect of some of these brain processes without any effective role. But although a simpler account is preferable, other things being equal, it is not so when such an account does violence to the complexities involved. If facts are complex we may have to prefer a more complex account however much simplicity is desirable on abstract grounds. I shall discuss later on why I believe that the influence of consciousness in initiating a particular brain process need not be construed as a complete break in the functioning of physical laws, although I admit that the mere entry of consciousness as an influencing factor into the field of human behaviour does mean that physico-chemical laws are not functioning here in precisely the way they function in the behaviour of unconscious matter, since in the latter case nothing non-physical is involved. When I say that they function differently I do not mean that there are different laws of physics and chemistry applicable to the human brain, only that some instances of the operation of these laws come to be active in the human brain partly through the influence of something non-physical such as consciousness.

The question for a theory of mind may be, (1) whether all human performances, including those which involve thought, can be explained in physical terms (meaning that it can be shown that such performances cannot take place without the occurrence of certain appropriate physical states) or, it may be as Armstrong asks, (2) can we give a complete account of man *purely* in physical terms? These two questions are different. For if the answer to (2) is 'yes' then as Armstrong puts it, '. . . a man is a physical object distinguished from other physical objects only by the special complexity of his physical organisation, and the special complexity of his physical capacities'. But the answer to (1) may be 'yes' and yet Armstrong's conclusion need not follow. For all human performances may be explained in physical terms, if 'explain' means what I have said above – particularly if we take the term 'performance' in a restricted way – and yet it may not be possible to

give a complete account of man *purely* in physical terms and man may be something more than a complex physical object. The objection to the materialist theory is an objection to saying 'yes' to the second question rather than the first. A psychophysical parallelist, for example, would agree that no matter what performance we take it is possible to give a complete translation of it into physical terms, that is, give an account of it in terms of happenings in the brain and the nervous system. But he would also claim that running parallel with the performances as pieces of overt behaviour are mental happenings which are not so translatable and he would say that a complete account of man must include both these orders of happenings. Thus according to this account performance as a piece of behaviour has a mental counterpart which is not physical and yet a performance can be translated into physical terms so long as we take 'performance' to mean behaviour minus conscious experience. That is, it claims that man is more than a performing animal; he has mental experiences and these are just as important as performances for understanding man as a total being. The parallelist may of course be wrong in saying that there are two different orders of happenings in the occurrence of all human behaviour but I am only mentioning him to show that to say that all performances can be explained in physical terms is not the same as to say that a complete account of man can be given purely in physical terms.

The materialist theory then has to claim not only that all human performances can be explained in physical terms – a claim most people would accept – but also that such an explanation is a complete account of man, that is, there is nothing more to a human being than happenings in the brain and the nervous system – a claim which some people find hard to accept. This theory can take two forms in trying to make good the claim: it can deny that there are mental experiences at all, as some materialists do, or it can admit that there are mental experiences which are existentially different from physical happenings but claim that these mental happenings are nothing but physical happenings even though they appear (illusorily?) to have a peculiar status of their own as felt experiences in which capacity they also appear to be non-identifiable with physical states or processes.

Armstrong admits to there being conscious experiences, he makes a distinction, however, between his materialism and epiphenomenalism of Huxley. According to Huxley we do have experiences which are in themselves something more than the workings of the brain but they are incapable of influencing the operation of the brain in any way. What

exactly is meant by saying that they are something more? Is it that experiences are there and are felt to be there by an experiencing subject who has no knowledge whatever of brain states? For if the experiences were existentially identical with brain states they could only be known as somehow qualifying the brain states, they could not be known apart. Does Armstrong deny that experiences are there and are known to be there apart from brain states? He does not, for he says that we have awareness of our internal states which is the same as having an experience. How then does his account differ from Huxley's who says that conscious experience is something more than a brain state? Does Huxley mean that experiences are not only there – and Armstrong also admits that they are there – but that they are non-physical in nature, while Armstrong thinks that they are physical? But what is meant by saying that a state of conscious awareness that p which one is experiencing is physical? And how does one decide whether such states are physical or non-physical once they are believed to be there as a result of certain brain states? For Huxley they are by-products of brain states, for Armstrong they are directly caused by brain states. But whether they are by-products or caused they are what they are and if they can emerge as non-material when arriving on the scene as a by-product they can equally do so when arriving as caused. And if they arrive as material when caused they can equally do so when arriving as by-products. I therefore do not see that there is any great difference between Huxley and Armstrong insofar as they both say that conscious experience is there and is produced by a brain state.

Armstrong no doubt uses the language of identity as well but whether conscious experience is 'identical' with a brain state or not, it is there and felt to be there as existentially distinct from a brain state – a brain state can be visually perceived, a state of conscious experience cannot be – and the mere assertion of identity does not take away the felt identity of a state of consciousness as experience. Is a state of consciousness capable of logically independent existence? Armstrong suggests that it does not really have separate identity if it cannot have logically independent existence. But why is it not capable of logically independent existence? It is a contingent fact that conscious experience is known (barring certain phenomena called 'extra-sensory') to happen consequent on certain brain states. Does this make it logically incapable of not being dependent on a brain state? I do not know how one proves this logical incapability, for to say that it could not happen without a brain state – and this does not follow from the fact that most human

experiences happen in the context of certain happenings in the brain – is already to assume the conclusion to be proved.

In Armstrong's materialist theory a mental state is defined as a state apt for causing certain sorts of behaviour. When defined in this way mental states are literally identical with brain states. We call certain states 'brain states' when they are looked at as physico-chemical functionings, we call the same states 'mental states' when these functionings are looked at as apt for the production of certain sorts of behaviour. The two different kinds of language, 'the brain' and 'the mind' are used with respect to the same existentially identical facts from the point of view of two different terms of reference. But people who object to the identity thesis do not define 'mental' as apt for causing a certain behaviour. So defined mental states do become identical with brain states, as mental here means $mental_1$. They include in it conscious experiencing of certain states of affairs as so and so, that is, $mental_2$. As soon as this is brought on the scene it can no longer be claimed that a mental state (which includes conscious experience) is existentially identical with a brain state, for conscious experiencing is felt to be an occurrence of its own kind quite unlike the occurrence that is a brain state. A man who is going through some conscious experience may look at some brain state which is supposed to represent it physically and find that he is being confronted with two disparate sorts of occurrences, although if we take a mental state to be a state apt for causing a certain behaviour there is no difficulty about identifying it with a brain state.

Armstrong of course includes a state of consciousness under the general terminology of mental states where they signify $mental_1$ and explains it away as yet another one of those states which are apt for causing certain behaviour. But it is not necessarily the case that a conscious state is a state apt for causing certain behaviour unlike a mental state which is so apt by definition. Experience may occur without leading one to undertake any activity at all. I may sit in a park and consciously enjoy the beauty of the day, which is to say experience that it is a beautiful day, without that having to produce in me any readiness for action of any particular sort. However, even if conscious experiencing is a state which is apt etc., it is still distinct from a $mental_1$ state which is nothing but a state apt for causing certain behaviour, say, my perception of the red signal as I drive along which is mental in the sense that it causes me to stop the car and yet not mental in the sense that I do not necessarily consciously experience that the signal is now red. A

mental state involving conscious experiencing cannot therefore be on a par with a mental state not involving conscious experiencing even if both are apt for causing certain behaviour. This is simply because in the former case it is possible to distinguish two different orders of facts present, experience and brain state, whereas in the latter, where conscious experiencing is not present, there is only the brain state and to call it a mental state is merely to point to certain logically distinguishable properties which we do not pick out when we think of a state as merely a physical state (meaning electrical discharge). Thus whether or not conscious experiencing is a $mental_1$ state, it is something more besides if a $mental_1$ state is a state which consists solely in being apt for causing certain behaviour. Armstrong also supports his case by using the language of cause – a conscious state is produced by a brain state. Now causal language does not involve the identity thesis except in the sense in which an effect may be said to be identical with its cause; but in this sense the whole world would be one identical unit. Armstrong must say that a state of conscious experiencing *is* a brain state and not just that it is the effect of a brain state, and I find this either false or meaningless.

According to Armstrong there are weighty empirical reasons against the dualist theory of interaction. It seems to me that reasons are not wholly empirical and amongst the reasons against dualism is the *a priori* doctrine about how all material systems can possibly work. For as Eccles points out the brain may be unique amongst all material systems wherein the pattern of discharge of neurones may be affected by consciousness. Thus all known material systems may function on a monistic principle while the human personality may be an exception to this rule. This would not necessarily mean however that discharges might occur in neurones which, according to Armstrong, could not be predicted by any physical law. This would have been the case if consciousness created certain physical states in the brain out of the blue, that is, if consciousness produced a state preparation for the emergence of which had already not been made in the physical organism in terms of antecedent states suitable for the production of the state in question. Whatever brain state occurs it occurs because antecedent states which make this state possible are already there in the organism. The influencing of a brain state may then simply mean that the organism, physically speaking, is in a state of readiness for more than one state to take place and the actual state that does occur is influenced in its occurrence – as against some other state the possibility for which is also

there and which could have occurred – by a state of consciousness. So some patterns of discharge in the brain may be said to be such that their actual occurrence – not the possibility of their occurrence as a pattern – cannot be uniquely predicted purely in physical terms, provided of course that the possibility of their occurrence is not inconsistent with the existing state of the organism. Thus if there be patterns of discharge in the brain which are influenced by consciousness, they do not so much contravene physical laws as complicate their functioning. For the activities of the brain are so complex – and man is an animal subject to diverse influences at the same time – that a present state of discharge of the brain may potentially contain more than one pattern of future discharge and consciousness may influence the actualising of one pattern rather than another, the possibility for the actualising of which is also potentially there inasmuch as the brain consists of a great many states and a great many connections between them.

Armstrong takes the interactionist doctrine to mean that a mental state is always interposed between two physical states, one which gives rise to a mental state and the other which follows it, whereby a gap is created between the two physical states and not only could the last in the series physical-mental-physical not be predicted from the first, but it could not be considered a consequence of the first according to known physical laws. I do not know if all interactionists do think in this way, but it plainly seems to me wrong. However, an interactionist need not claim that interaction is involved in every single instance of human behaviour, nor need he claim that in the instances where interaction of a sort does take place a complete break with physical causality is involved. This last point is recognised even by Armstrong for, as the following quotation shows, he finds an interactionist saying that an action may occur which is not *solely* determined by a physical state. Not being solely determined by a physical state does not however imply that it is wholly undetermined by such a state. To quote,

> Let us now consider a situation where a physical stimulus of some sort, say the sounds of a human voice, brings about certain mental events, say perceptions and thoughts. On the 'way up' there must be a last physical event in the brain before the mental events ensue. The mental events must then bring about a first physical event in the brain on the 'way down'. So this second event will have to be an event that is not solely determined by the previous physical state of the brain and physical environment. This means that there will be as it were a 'gap' between the state of the brain before the mental event has had its effect and the state of the brain after the mental event has had its effect. The transition between the two states will not take place solely as a result of the physical workings of the brain and physical environment.[1]

Let us take a concrete example to see what is involved. I am passing down the corridor when I hear a human voice. I perceive the voice as that of a friend of mine and turn back to greet him. To say that I perceive the voice as that of a friend of mine means that I am ready for a certain sort of behaviour that fits the fact that the voice is that of a friend. Here no specifically 'mental' influence is involved and the electric discharge in certain neurones caused by hearing the voice leads to the electric discharge involved in my turning back. But suppose that I perceive the voice as of someone about whom I am in two minds as to whether to greet or not. I think I ought to say hello, for not to do this is unfriendly and I do not want to appear unfriendly, and yet I also want to avoid him if I can at this particular juncture as I fear he may remind me of a promise I have not yet kept. I continue passing down the corridor uncertain for a time as to what to do and then I decide to turn back saying to myself 'Well, I might as well face it'. In this case it seems to me that the discharge in neurones involved in hearing the voice does not uniquely determine a particular pattern of discharge, either walking on or turning back, potentiality for both of which is there in the functioning of the system of neurones that is my physical organism, and my conscious decision to do one or the other activates one particular pattern in preference to the other which could also be activated as far as the state of the physical organism is concerned. What comes about is thus not determined solely in physical terms and yet it will be wrong to say that my action was not at all determined by physical causation or that it involved a gap in the functioning of physical nature such that a physical state comes about which can in no way be accounted for in terms of antecedent physical states. Even if my turning back is a result of decision a neurologist looking at the two patterns of neurone discharge, hearing the voice and turning back, can connect them in physical terms. Equally, had I decided to walk on he could still connect the two patterns of discharge, hearing the voice and walking on in physical terms, and there has been no break in physical functioning of

[1] Ibid, 32.

a such nature that we have to say that an action that followed is absolutely unpredictable on the basis of antecedent physical states (even though not uniquely predictable).

Armstrong would not find this acceptable not because, so it seems to me, it is impossible for a certain state of the brain at a certain time to hold the possibility for alternative patterns of discharge, the actual pattern that follows being one that comes to be activated as a result of some conscious activity on the part of the agent, but because he holds the monistic doctrine that there is nothing but physico-chemical energy at work in the universe and each state of discharge of such energy uniquely determines the succeeding state. The account I have given of a conscious process – which is not thought to be a bit of discharge of physical energy – influencing the occurrence of a certain pattern of neurone discharge, is thought by the materialists to be mysterious and hence suspect. But why should it be thought mysterious? As Armstrong himself admits there is no logical implausibility in the thought that a non-material process influences a material one and this may happen only in the functioning of a human personality and nowhere else in the universe. The implausibility then must be shown to be empirical, that is, neurophysiology must show that it is, empirically speaking, impossible for the brain to function in this way (to show that matter does not function in this way elsewhere is not good enough for the purpose). This has not been done nor can it be done merely by showing that unique causal determination of a purely physical nature exists in the brain. For nobody denies that it does exist and it may exist to a far greater extent than has hitherto been thought. The question is whether this is all that there is in the functioning of a human personality. The implausibility is then doctrinal which assumes, on the strength of the success of physical sciences to explain lots of things which had at one time been thought to involve non-physical causes, that there is nothing non-physical in the universe. But although a great deal that once 'belonged' to the non-physical is now rightly believed to be physical the very existence of the non-physical need not be questioned, if such a concept is needed to explain experience. It may well be that it is only a material system of the degree of complexity that the human brain is which allows for the influencing of material functioning by a non-material principle. And the reason why we bring in the concept of a non-material principle is that we cannot otherwise satisfactorily explain the fact of human consciousness and freedom. This will explain why there is just this one thing, as the materialists complain, where an exception has to be made when the rest of the things of the world can be explained purely in physical terms.

Now when the materialist theory of mind denies that privately experienced inner states called 'mental states' occur in a separate sphere from neurophysiological states whenever we have sensations, feelings etc., this denial may take two forms. It may be a denial that (1) there are mental states which occur as psychical counterparts of physical states whenever we have sensations or (2) it may be a denial that there is such a thing as experience which is private to oneself. Materialists often do not make any distinction between these two types of denial and they believe that when they have shown the so-called counterpart mental states to be non-existent they have also shown that there is no such thing as private experience. (And even if there is experience, as some materialists allow, it is held to present no problem since it is treated under the general account given of mental states when such states are identified with brain states.) This does not seem to me the case. My purpose in separating the two aspects of this denial, or what I think are two separate aspects, is this. I agree with the materialist denial (1) but not with denial (2). That is, I agree with the materialist that we do not need mental states or processes which are things over and above neurophysiological states and processes in order to explain human performances including performances which involve speech, overt or covert. I also sympathise with the behaviourist contention that when we try to pin down these so-called mental states occurring in a sphere of their own they seem to evaporate and we are left with nothing but a description of our responses which are neurophysiological in nature. If someone asks me, 'what is your mental state of seeing red like?' I can only say that I am having an experience of something which I see as red and not what seeing red is like as a mental state. There are not then two different kinds of seeing involved, mental seeing which is private and is not produced by a brain process and physical seeing which is a brain process, even in a case where I am having an experience that I am seeing red.

So far as seeing something as red is concerned, it is the same seeing caused by a brain process whether I am conscious or unconscious of the fact that I am seeing red, and when I become conscious of seeing red, no mental state of seeing red which is additional to physical seeing appears on the scene. All that happens when I have an experience is that I become consciously aware of the fact that something is red and also that I happen to be seeing it. This conscious experience is usually described

as a mental state of seeing red as distinct from a physical state involved in seeing. But this terminology is confusing, for when I become conscious of seeing red no additional seeing takes place and the same seeing red is involved as was there when I saw red but without any conscious experience of that fact. The only difference between the two cases lies in the presence or absence of consciousness that *p*, not in the presence or absence of a state of mental seeing which is an internal duplication of the physical seeing without that internal state being caused by a brain process. Whether I have conscious experience or not of what I see, the evidence that I have seen something lies in my response, such as my ability to describe it. In this sense seeing is public whether attended by consciousness or not and it does not consist in an internal state private to myself of which nobody has any knowledge.

But conscious awareness that seeing red does involve something private although seeing red does not, and this is a state of experiencing that conscious awareness stands for. We should not call it a *mental state of seeing red*, for the experience is not experience of seeing red, it is experience *that* seeing red. But experience that seeing red is a state of conscious awareness and in that sense it can be called a mental state. Nevertheless as a state it is not a mental duplicate of physical seeing red, it is only a state of being consciously aware that red. It is indeed a state of consciousness but it is not a conscious state representing seeing red, it only represents awareness of seeing red. Thus it is not necessary in order that we may be said to see red that there be a mental or conscious state of seeing red over and above a physical state and there are no conscious or mental states representing our thoughts, sensations etc. But there can be a conscious or mental state to the effect that thought *x* or sensation *y* has taken place. And this conscious or mental state is the same *qua* its conscious or mental characteristics whether the state is that thought *x* or sensation *y*, and I can only distinguish a conscious state that thought *x* from a conscious state that sensation *y* in terms of the thought or sensation involved, not in terms of consciousness (for consciousness as a function is the same no matter where it is directed). This is why states of consciousness have no identifiable features which mark them off purely in terms of consciousness minus our responses towards which they are directed. Nevertheless there is such a thing as a state of conscious awareness that *p* and of this we have experience. In other words, there are no states of consciousness to be called *x*, *y*, *z* which are distinguishable in terms of consciousness and which duplicate physical activities on an

internal screen, but there is a state of being conscious that *p* which can stand for a sensation, a feeling or any $mental_1$ state.

Thus to say that seeing red does not need an internal mental state is not necessarily to say that there is no such thing as private experience. It is perfectly possible for me to see red as I do when I stop my car at the red signal without going through any private experience that I am seeing red. Here I must be said to have seen red for I would hardly have stopped otherwise, but clearly nothing other than a brain state is involved. But this is not an instance of human behaviour or functioning which would be offered as an example of private experience. But suppose that at the red signal I not only stop my car but become consciously aware that the signal is red, possibly because I am in a hurry and therefore more likely to become conscious of obstructions. Here something has happened which is private to myself over and above a publicly observable response such as that of stopping the car. A fellow motorist watching me knows that I have seen red but he does not necessarily know that I am having a conscious experience that the signal is now red. This conscious experience is not necessary for seeing red or for stopping the car but it may occur on the occasion of seeing red. And when it occurs it cannot be explained purely in terms of states that are involved in seeing red, as something more is clearly involved.

As I have already observed, I agree with what the materialist has to say about $mental_1$ states, that these states, as are involved in activities like seeing red, are not really states in a separate psychical sphere but brain states themselves considered under the aspect of certain logical relationships. But I do want to insist that a state of conscious awareness that *p* cannot be said not to occur as something distinct from seeing red when that consists in merely responding in such a way that fits the description 'seeing'. This state cannot be studied via performances as, contrary to Armstrong, it is not a state apt for causing certain behaviour, in the way the mental state of seeing red is. It just consists of an experience and does not necessarily have any performatory significance. I stop my car when I see red, I do not do anything extra when I have a private experience that the signal is now red, nor do I perform the act of stopping the car with any greater efficiency because I am having this experience. Of course it is only the man having this experience who can vouch for it by saying, for instance, that he is having it, and there is no publicly observable criterion, such as is available in the case of seeing red, by which we can test whether a private experience has taken place

or not. This is why it is suspect in a behaviourist-materialist account of man.

AN EXAMINATION OF SOME MATERIALIST-BEHAVIOURIST ARGUMENTS AGAINST EXPERIENCE

One of the least important (but apparently to the materialist a very weighty) argument in favour of saying that there are no mental phenomena, is, as said by Smart,

> Mental phenomena, as the Cartesian conceives them are 'queer' entities the existence and operations of which violate the principles of science, both the principles discovered and established by science, and the principles which specify the nature of legitimate scientific enquiry itself, principles of evidence and intelligibility, for example.[1]

Note the use of the term 'queer' here; it is being used in respect of the mental because that does not fit into the scheme of science. This is perfectly all right as long as we understand that the term is being used in a technical sense introduced by definition. 'Any object which cannot be scientifically studied is queer' is then a tautology from which it does not follow that things which are queer in this sense cannot possibly be there. But this is the conclusion that is implied through a surreptitious shift from this tautologous use of 'queer' to its ordinary use where whatever is termed 'queer' becomes highly suspicious.

Mental phenomena are said to be queer because they violate scientific principles of evidence and intelligibility. This is somewhat vague, for are evidence and intelligibility that are missing scientific evidence and intelligibility (e.g. amenability to experiment etc.) which of course cannot be possessed by 'queer' entities by definition, or are they just evidence and intelligibility suitable to the phenomenon in question – these being principles that science advocates should guide our study of whatever we are trying to understand? For it does not seem to me that whatever is referred to as 'mental' (here of course mental means conscious) as distinct from physical, physiological etc. lacks evidence and intelligibility of a sort that is suitable to itself. The sort in question does not fit into the scientific scheme but this may not be a disqualification if all evidence and intelligibility are not necessarily of the sort which is defined by the scientific technique, After all, 'mental' or

[1] *The Philosophy of Mind*, ed. V. C. Chappell (Prentice-Hall, Englewood Cliffs, N. J. 1962).

conscious' is not a technical term introduced by philosophers, they are in common use and perfectly well understood by ordinary people who do not doubt that they have evidence for what this term means in an immediate experience which consists of a conscious awareness of something. I do not wish to say that ordinary notions may not be subject to confusion and some of the uses of the term 'mental' or 'conscious', not only in common parlance, but also in philosophical literature may well be confused. Nevertheless, the claim that the referents of these terms lack all evidence and intelligibility, does not seem to me substantiated by the ability of ordinary people to communicate to others what they mean when they use these terms. The reason for this is that ordinary people do not suffer from the scientific bias and consequently do not demand that all evidence and intelligibility must be of the sort which go with the use of the scientific technique.

I have said above that ordinary people believe that they have evidence of the occurrence of the mental (meaning conscious) in immediate or direct experience. But of course evidence of this sort is not involved in scientific verification and there are philosophers who believe that the term itself is vacuous and has no real referent apart from an illusory one 'fostered by current prejudices'. This is the position taken by B. A. Farrell in his article 'Experience',[1] where he criticises physiologists and psychologists who believe that a full description of the mechanism involved in the activity of the brain and the nervous system would still leave a gap so far as a proper understanding of human beings is concerned. He quotes Adrian as saying 'the human mind comes in somewhere in the chain of causation between the physical events in the sensory and motor pathways.'[2] As to the nature of this connection Adrian accepts Kenneth Craik's suggestion that 'consciousness comes in at certain points in the process of neural transmission where the physiological patterns have a particular kind of definiteness . . . on this hypothesis we could tell what someone is thinking of if we could watch his brain at work, for we should see how one pattern after another acquired the necessary brilliance and definition.'[3]

This appears to Farrell as nothing but a manifestation of the occupational disease of traditional dualism which bars the way to a complete account in purely neurological terms of what happens when we think,

[1] *The Philosophy of Mind*, ed. V. C. Chappell.
[2] Ibid., 23.
[3] Ibid., 23.

recognise, remember and see things – an account which should ignore the mind and all intervening mental wants whatever, as these throw no light on what happens when we think etc. Now I think Farrell is right if Adrian and Kenneth Craik are saying that whenever we think, meaning use words intelligently, aloud or in silent speech, consciousness comes in and influences the process of neural transmission. I agree with Farrell that such a supposition may be a bar to a complete account in purely neurophysiological terms of what happens when we think. I do not think that consciousness (in the sense of $mental_2$) is an agency which conducts thinking through influencing the neural organism. I also think that thinking, insofar as by this is meant intelligent use of words, aloud or silently to oneself (that is, $thinking_1$), can very well be understood as an interconnected neurological event by itself. But there are times when conscious awareness of what we are thinking of does influence the subsequent use of words, the use of these words of course being neurologically underwritten in the organism. I disagree with Farrell because he appears to write this possibility off on the ground that thinking, that is, thinking without an attendant conscious experience of what we are thinking of – an experience which may come to influence the subsequent use of words – can very well be understood as a neurological event without any psychic trappings.

It is a mistake to think, says Farrell, that a physiological and neural account of what happens when we think leaves out certain essential parts or aspects of the total process – parts or aspects which are expressed by saying that one is having experiences and which the psychologists can investigate. This very well fits the case of $thinking_1$ but not $thinking_2$ (which Farrell does not consider at all) where consciousness does come in and where something is left out if a philosopher says that there is no such thing as conscious awareness, and that this never influences what goes on in the neurological organism.

Farrell holds that to find out what happens when X has sensations, all that is necessary is to subject him to stimuli that are likely to affect X's sense organs, vary them in various ways and observe and measure X's responses and discriminations. Everything that the psychologist wants to know about sensations can be discovered by making X behave like a dog, it is not even necessary to make X talk. Of course the psychologist himself can be the subject as well as the observer. But he does not thereby obtain anything new that he could not in principle obtain from the observation of X alone and no new concepts other

than behavioural ones are involved. But some people feel that to study behaviour is not to study sensation, the raw feel of experience, as Tolman named it. But even according to Tolman one does not know what to do with these raw feels; they cannot be studied scientifically – they can either be assumed to correlate consistently with our responses and response readinesses so that so far as X and Y behave alike they have the same sort of experience or they may be the way physical realities are intrinsically, so that experienced qualities may constitute the intrinsic nature of a nervous process. This appears to Farrell metaphysical nonsense and he declares that were our knowledge complete, so that we had a perfect correlation between sensation and neural processes, we should then identify the two. That is, he appears to hold that there is no such thing as raw feel of sensation which is privately experienced over and above publicly verifiable responses and response readinesses, or in other words, there is no experience attached to sensations, it is all a matter of behaving in a certain way.

Now the term 'sensation' is ambiguous. It may denote just what Farrell says it does, responses and response readinesses. I have a sensation of pain when I touch, say, a hot plate, and the evidence that I have a sensation of pain lies in my response of withdrawing the hand. This is $sensation_1$ and it does not involve any raw feel of experience. But I may become consciously aware of the fact that I am having a certain sensation and this awareness gives my sensation an experienced or felt quality, that of pain, identified as such and such. This is $sensation_2$, and often it is $sensation_2$ which is implied when one talks of sensation of pain and which suggests something mental. But $sensation_1$ is mental too, although it is not conscious like $sensation_2$. I agree with Farrell, that $sensation_1$, the withdrawal of my hand in response to pain, does not depend on $sensation_2$ nor does it necessarily involve it. But $sensation_2$ does take place, it does have the raw feel of experience and it is only in respect of $sensation_2$ that people say that something is left out if you reduce it to responses and response readinesses.

I agree that $sensation_2$ is not a mental counterpart of $sensation_1$, it is a state of awareness that $sensation_1$, which confers on $sensation_1$ an experienced quality. Apparently it is only in $sensation_1$ that Farrell is interested and of course its study does not need an understanding of $sensation_2$. But for someone who is interested in the fact of $sensation_2$ when it does take place it will not do to say that all that is involved is responses and response readinesses. Now since $sensation_2$ is not a mental counterpart of $sensation_1$ which occurs whenever $sensation_1$ occurs,

it may well be that nothing is left out when one gives an account of $sensation_1$ in terms of responses. But if $sensation_2$ does attend $sensation_1$ then it will not do to say that experience is not left out in such an account. Farrell does not appear to make this distinction and it seems that he even questions the very existence of $sensation_2$. Attacking the idea that experience is a component of sensation he says that what we are supposed to leave out, that is, experience, when we identify X's sensations with his differential responses and response readinesses, is not at all like what may be left out when one says 'If you merely consider what Y says and does, you leave out what he really feels behind that inscrutable face of his'. This says that if you rely on the publicly observable behaviour of the person you will be ignoring his private experience and you may be wrong in the guesses you make about it. Here Y can tell us what is left out if anything is left out but X cannot. He may give us further reports but this will be about his discriminations etc. So we are still left with behaviour while what we want is experience. But X's experience behind his discriminations cannot be got at or talked about, it is featureless. This is not true of Y's experience behind his inscrutable face. X's experience is featureless because there is nothing about it that X can discriminate. Further, when we say 'X's experience is featureless' it looks as if something is being said about his experience. However nothing is being said about the supposed raw feel nature of his experience, whether it is a thing, process, event or state of affairs. This indescribability of 'experience' is different from, say, the indescribability of the beauty of the Alps. The indescribability of experience is logical, since there is nothing to describe; not so the beauty of the Alps.

I am not quite clear as to what all this is supposed to prove. It seems that Farrell is saying that there is no such thing as a conscious experience of seeing red, that is, $sensation_2$, as this cannot be described except in terms of seeing red, that is, $sensation_1$. Experience itself, which is supposed to be something more than $sensation_1$, has no feature. Presumably then it consists in nothing other than $sensation_1$. It seems to me natural however that experience by itself has no describable feature. For experience of seeing red is an experience that one is seeing red or that what one is seeing is red, and so naturally if asked to describe it one does so in terms of seeing red. The experience is not experience of seeing red which is a mental counterpart of physical seeing (brain process), it is experience that seeing red. Experience means a state of conscious awareness that such and such is the case and to say that it has no feature means that it cannot itself, *qua* experience, be described as red, loud or as any other quality applicable to $sensation_1$. But why should this be a disqualification? A state of conscious experience *qua* consciousness is simply not one of those things that can have a feature of its own but we can hardly say on this account that we have never experienced what it is to be in a state of conscious awareness, nor can we say that we do not know what it is like simply because we cannot describe it. This indescribability is not peculiar to consciousness. Suppose that I have correctly identified a colour as deep purple. Then someone asks me to describe it, for all I have done so far is to name the colour. But deep purple has no further feature by which it can be described. This does not arouse suspicion because colour is a public quality and it can be ostensively defined when words fail, whereas my consciousness is private. Nevertheless, non-philosophers seem to have no difficulty in understanding that we all have private experience. Also, even if my private experience is not publicly shareable, it is publicly referable besides being privately identifiable by me. These are the reasons why we do not think that it lacks all evidence and intelligibility.

Farrell continues with his criticism of the concept of experience. Some people, he says, may insist 'But we just do know what it is like to have the experience of looking at a red disc and we just do know that this differs from the experience of looking at a green one'. But to know just what it is like to have the experience of looking at a red disc means no more than that we have the capacity to respond symbolically, unlike the rat which can be conditioned to discriminate between red and green, to the pattern of stimulation produced by the response of looking. When therefore I say 'I just know what it is like to have the experience of looking at a red disc' I am saying that I have the capacity to discriminate my own responses in the same sort of way as I discriminate the features of the red disc when I say of it, for example, 'Oh, it is bright, uniformly red' and so on.

Farrell thus makes no distinction between the two levels of discrimination, (1) responding differently to red and green, as both rats and humans do, and (2) being consciously aware, in the context of responding differently, that the two stimuli are different. This second level may be expressed in symbols as in 'I just know that the experience of red is different from that of green'. This is described by Farrell as, 'responding symbolically and discriminating my responses'. Farrell thus identifies having conscious experience with articulation and use

of symbols. This does not seem to me justified. For it is perfectly possible to use symbols without any conscious awareness being involved that one is discriminating such and such qualities, also, one may have conscious experience and use no symbols (except that some form of internal speech is involved in conscious awareness). These two capacities for conscious experience and for use of speech, are very closely related in human functioning. It is only human beings who are capable of becoming consciously aware of distinctions. And the human capacity to use symbols is linked with the capacity for conscious awareness, since symbols are used to codify distinctions and sometimes distinctions referred to by language are such that one needs to be consciously aware of them in order to grasp their full significance. Rats do not use language and their incapacity partly, at any rate, is linked with their incapacity for being consciously aware that so and so. The level where conscious experience that certain discriminations are being made may be involved and where in addition language can be used in respect of discriminations made is then different from the level where we, rat-like, just respond differently to red and green. This distinction cannot be further described and one may just have to insist that one knows the distinction by knowing what is involved, as one does when one claims, 'I just know the distinction between red and green'.

I am of course not saying that when the rat responds differently to red and green it does not know, is not at all aware of, the distinction between red and green. But there is a difference between knowing the distinction between red and green, which is shown in appropriate responses, and knowing *that* there is a distinction between red and green in the context of one's being consciously aware that this distinction is there, and the difference lies precisely in the presence or absence of conscious awareness.

Some people may object that rats must have experience of the distinction between red and green in order to respond to them differently. If so, we have to call this experience$_1$ and conscious awareness that *p*, experience$_2$. Experience$_1$ is experience of *p* as shown in correct response to *p* and its sole reality lies in this response; experience$_2$ is experience that *p* which is no doubt shown in correct response to *p*, but it is not exhausted in this response. It has a felt quality of its own and it does not consist merely in a readiness for response.

Farrell does not see any difference between the two levels of awareness and he believes that the capacity to respond differently to red and green and the capacity to respond to this capacity symbolically, are exactly of the same order. But even if the capacity to use symbols, which consists of certain electrical discharges in nerve cells, is of the same order as the capacity to respond differently to red and green, it would not show that conscious awareness that red, which consists in more than a capacity to use symbols, is the same too. However, it seems to me that the capacity to use symbols is closely linked with the capacity to have experience$_2$. I do not mean by this that every time a man uses symbols he is consciously aware of the distinctions referred to by these symbols. Habitual use of symbols does not need any such conscious awareness. All I am saying is that the very device of using symbols to mark distinctions can come into existence on a level of response where conscious awareness of such distinctions is possible. To say then that the only difference between the responses of rats and human beings is that human beings are capable of responding symbolically while rats are not, is to say that rats are not capable of experience$_2$.

As Farrell does not see this distinction he remarks,

> If we overlook this capacity to react symbolically to one's own reactions, or if we say we are doing 'introspection' and do not inquire what actually happens when we do so, then we are liable to assert with a flourish and an air of importance that 'we just know how looking at a red disc differs from looking at a green one'. In other words, we are liable to mistake features of our responses to the disc for some indescribable and ineffable property of the experience.[1]

Thus Farrell tries to reduce experience$_2$ into something to which the behaviourist criterion applies; it consists in our ability to respond symbolically. But although capacity for experience$_2$ is closely linked with capacity for symbolic response it will not do to say that the two are the same. There are occasions when I do have experience$_2$ but I am not responding to anything symbolically – that would not show, at least to myself, that I am not having an experience, although other people may have no occasion for saying that I am having it. Also, I can use symbols without going through any conscious experience. Experience$_2$ is generally described as private, but it is publicly referable and some of the symbols that we use, such as, 'I know that *x*' are a correct clue as to its existence. But if I do not talk about it when I go through an experience it remains private and its only proof lies in

[1] Ibid., 34.

itself. Also, experience$_2$ takes place in the context of experience$_1$ so a description of experience$_1$ coupled with my use of a symbol in respect of this experience may furnish an indirect reference to experience$_2$. But neither experience$_1$ nor the use of a symbol to name this experience guarantees the presence of experience$_2$. Since this experience$_2$ is not directly revealed in a behaviourist study of first order response, including use of symbols, we say that we know about this introspectively, that is, we say that we know the distinction between red and green by knowing this distinction in the context of being consciously aware that red and green rather than behaviourally, that is, by studying our first order responses.

As I have already remarked I am not claiming that whenever we respond differently to red and green, a parallel internal process of consciously knowing that we know the difference between red and green is going on, which must be studied introspectively. However, consider this case. Someone says to me 'pass that red book over to me please', and I do so. Suppose then he asks 'how do you know that you know this is red?' I am hardly likely to reply, 'didn't you see that I responded correctly to your request by not giving you a green one?' I am more likely to say, 'well, I just know that I know'. This is expressed in more technical language by saying 'I know it introspectively'. This is hardly an answer to a normal 'how'. 'How do you know he is rich?' 'I have looked at his recent bank-balance' satisfies the query. 'How do you know that you know red from green?' 'I know it introspectively' seems disappointing, for knowing something introspectively simply means knowing it just by becoming consciously aware of knowing it. But this is not superfluous in a complete understanding of human beings. The rat too knows the distinction between red and green but it does not consciously know that it knows, that is, it knows the distinction behaviourally but not introspectively.

I would agree with Farrell's denial of experience, if 'experience' is to mean a psychical event with describable and identifiable features of its own (in the way brightness is an identifiable feature of red). Experience of red, which really means experience that red, is not describable except in terms of red. We may call this experience a psychical event but it is not a psychical event of the nature of experiencing of red, which can be described and identified by certain features of which it is an experience. Experience$_2$ is an experience that something is the case rather than a mental event which can be described as an experiencing of that something. Farrell, however, rejects experience$_2$ and suggests that it consists in nothing but a capacity to use symbols. So experience$_2$ is identified with experience$_1$ and it gets reduced to response and response readiness. When I object to this reduction I do not imply that experience$_2$ takes place independently of all neurophysiological relationships. It is only a human being with a very complex and highly organised brain who can have experience$_2$. But whatever the relationship between this experience and neurophysiological processes it can hardly be that of existential identity. Nor can it, according to me, be physical causation. To have experience$_2$ is not only to be consciously aware that *p*, it can also involve a self-conscious awareness that one is having an experience that *p*, that is, it involves in Farrell's words, discriminating one's own response to stimuli, and not merely features that belong to the object experienced.

Thinking, says Farrell, is intelligent use of words and not something else besides of a psychical nature. Suppose that X is thinking to himself about *y*. What is going on is silent speech and exactly the same words are there that would have been involved had he been talking to somebody else about *y* instead of silently to himself and nothing extra is involved. But the absence of this extra something does not make his talk to himself about *y* thoughtless. Nevertheless, it seems to me that the claim that there never is anything involved in a total process or event called 'thinking', which can be called experience, is mistaken. There may not be an experience called 'thinking', so that if we know that X is thinking of *y* we could say that X is having an experience *z* to distinguish it from the words which X uses in his thought of *y* – words which may be looked upon as responses. But there can still be a conscious awareness on the part of X that he is thinking of *y*, and this awareness confers on the thought itself a felt quality, which it did not have before when it consisted of use of words and nothing else besides. However, this felt quality does not change the character of thought which still consists in intelligent use of words. The describable features of this experience are naturally those of the thought of *y* rather than anything that belongs intrinsically to consciousness itself. In other words, conscious awareness that thought has no features *qua* consciousness by which it can be distinguished from conscious awareness that pain or anything else. Consciousness that thought differs from consciousness that pain by virtue of the difference of thought from pain as responses, rather than of two different characteristics of consciousness *qua* consciousness, which belong to these two conscious states. What I have said does not necessarily imply that there is any such thing as consciousness

which is not consciousness of anything and which can be inspected as just consciousness. But consciousness can be thought of *qua* consciousness and not *qua* what one is conscious of and in this sense can be referred to in a topic-free way.

Things then can be mental without having peculiarly mental features over and above behaviourally determinable features. If so, the absence of describable features does not necessarily prove the non-existence of what is called mental when mental means conscious experience. What about animals then? Can anything concerning animals be described as mental since they are not consciously aware that *p* even when it can be said with perfect propriety, when we apply a behaviourist criterion, that they know *p*? It depends on what meaning we attach to the term 'mental'. Mental_1 is used to designate certain features that characterise behaviour like intelligence, capacity to vary one's responses, adaptability to new situations and so on. Animals certainly do exhibit these and so they can be said to possess mental attributes. Animals, of course, do not have the capacity for speech, but speech as an intelligent response is a further development of the same sort of capacities that are being called mental_1. Mental_1 does not need, as an explanation of the capacity involved, occurrences other than neurological and behavioural ones, for here conscious awareness that *p*, which is, existentially considered, obviously something more than a neurological or behavioural occurrence, is absent. Human beings have mental characteristics of this sort as well as animals, the only difference being that the human capacities are much more highly developed. Mental_1 thus signifies any discriminative response or capacity for such response, and it includes highly intelligent response to stimulus where both the stimulus and the response are complex, involving hierarchically ordered performances, the lower grades of which feed back the higher, leading ultimately to a purposeful adjustment of actions to complicated requirements of situations. All this may happen entirely through the law-governed connections of neurophysiology without the intervention of consciousness.

Some philosophers then find that consciousness is not strictly necessary for an explanation of human behaviour and this seems to me right as far as it goes. If we take a human being as a performing animal of high intelligence who can further his own goals, as set by the environment via the requirements of his inherited constitution, by responding in a highly complicated manner to the demands made by the situation in which it is placed, insofar as these impinge on his goals, it is perfectly possible to think that the neurophysiological mechanism is enough to explain his activities, some of which would be described as intelligent, volitional, emotional etc. according to objective behavioural criteria. That is, given the brain and the nervous system all the complicated responses that human beings make to their environment could all be carried out even if they were not capable of experience_2. But experience_2 makes at least one significant difference apart from the difference that the fact of conscious awareness itself makes. It is because human beings can be consciously aware of the distinctions that they make, that they can come to have an understanding of alternative goals and alternative possibilities of action, between which a choice has sometimes to be made by a conscious decision. This capacity for decision, as well as consciousness itself on which it is dependent, cannot be explained, as far as I can see, as functions of neurophysiology, yet these alter the character of human response when that is thought of not merely as an overt performance but as a total reaction that includes deliberation, decision etc. as well as a conscious awareness of oneself as having a certain kind of experience. I shall deal with this question more fully in the section on freedom, but suffice it to say here that certain types of human reactions cannot be satisfactorily understood except through the use of the concept of experience_2.

Experience_2 introduces privacy. When I think, my thought consists of the words that I use silently to myself and it can be made public by uttering the words aloud, but not so my conscious awareness that *p*. Of course I can say aloud 'I am aware that I am thinking', and this makes other people know that I am having an experience. So despite privacy experiences are publicly referable. However, there remains a difference between the way they know my thought and the way they know my experience. When my thoughts are uttered they take on a character which is entirely public and as such shareable by all who understand the words used. But when other people know that I am having an experience through my saying so, they do not participate in the felt quality of my experience as I do through having it. So privacy of experience does not mean that nobody knows that others have experience, it only means that other people cannot know the felt quality of my experience as I am aware of it, although they know about the felt quality of experience through having felt it themselves. However, this inability to share in the privately felt quality of an experience does not involve inability to share what these experiences are about, for this we can know from behavioural criteria including verbal

utterances, and there is a great deal to be said in favour of the use of behaviourist criteria in studying experience$_1$ where intrusion of irreducible mental concepts supposed to bring such experience about may be an impediment to further progress. As behaviourism tries to give an explanation of human behaviour in scientific terms, in the form 'this happening that happens', and as there is nothing to experience$_2$ but the felt quality of conscious awareness without any performatory counterpart, which makes it unlike a piece of behaviour, it is perfectly legitimate for it to leave out consciousness or experience. But to deny that consciousness or experience happens at all is another matter. The behaviourist assumption that human beings do not have experiences private to themselves seems to me obviously wrong, even though to study a mental$_1$ occurrence – the only type of occurrence it is interested in – is not to study a private occurrence, and even though there is a sense in which private occurrences cannot be studied.

Behaviourism rightly insists that the only fruitful approach to the study of subjective or mentalistic concepts is in intersubjective terms. Of course what is being studied here is mental$_1$ and not mental$_2$, that is, intelligent response to *p* rather than conscious awareness that *p*. But if we are going to study thought as a human reaction or response to one's situation what we need to study, as the behaviourists insist, is the response itself, not a private awareness that thought, because the occurrence of this private awareness does not throw any light on the nature of thought as a response (by and large) nor can it be studied in intersubjective terms. Let us take for example thought which consists in, say, solving a problem in a jig-saw puzzle. How one thinks the puzzle out or what one's thought about a move consists of, can only be studied by an intersubjective process of observation of tentative attempts, actual moves made etc. – not by introspection. The only thing that can be studied about thought is thought$_1$ – thought that consists in certain moves or in the use of certain words – and the only ways of studying it are behavioural and neurological ones. It is only thought$_2$ that it is impossible to understand without the concept of consciousness, for thought$_2$ consists precisely in a conscious awareness that thought$_1$. So the behaviourist rightly concentrates on thought$_1$ which consists in behaviour involving speech and action, and leaves out raw feel of experience$_2$ which has nothing to say about behaviour, actual or possible. The objection then seems to me well taken that a psychology which wants to advance our knowledge of human functioning, including those called mental (mental$_1$), should have nothing

to do with the phenomenal field of raw feel and should confine itself to expressive behaviour and dispositional tendencies manifested therein. Even if mental$_2$ influences the course of the mental$_1$ in some fashion, in certain cases, mental$_1$ still remains expressive behaviour and dispositional tendency and it can still be studied according to behaviourist criteria. Only, if mental$_2$ happens and it does influence, to whatever degree, some happenings in the sphere of mental$_1$, to deny these facts is to give a false picture of what it means to offer a complete account of man as behaviourism claims it does. However, the study of mental$_2$ and its influence on behaviour does not belong to behaviourism; perhaps it belongs, if anywhere, to literature.

There seems to be a good case then for saying that experience$_1$ can be reduced to physical and physiological functioning. Experience$_1$ consists in intelligent functioning, which plainly appears to be a function of the brain and the higher the complexity of matter and form in the brain the greater the intelligence. It is then perfectly feasible to study intelligence and intelligent performance in neurophysiological terms and it may one day be possible to find out the precise neurological factors of which given degrees of intelligence or given types of intelligent functioning are functions. Translatability into neurology, however, seems possible only so long as we leave experience$_2$, the raw feel of awareness, out of our account of man, for then what is involved is a functioning or some sort of dispositional tendency or readiness for such functioning and there is no difficulty in thinking of it as a physical state. But directly we bring in experience$_2$, the raw feel of awareness, we are confronted with two things, at least on the phenomenological or existential level. Suppose that I could see the state of my brain with the help of a periscope hitched up to a microscope whilst I am having a raw feel experience, I would still think that I am being confronted with two different happenings, and I would not doubt that the orders of happenings involved are distinct. For I can perceive the brain state with my senses just as I can perceive the movements of my limbs when I perform an action or hear words when I utter them, but the order of happening involved in raw feel is such that it can only be had through conscious awareness and never perceived whether by me or by anyone else. Since it cannot be perceived either directly or indirectly through its effects it cannot be said to be a physical thing, explainable in terms of physico-chemical laws. It is thus not reducible to neurophysiology, at least not in the same way as performance, and not without the help of a theory such as materialism.

Of course things which are existentially different may yet be said to be identical when they can be described in indistinguishable terms, as two mass-produced coats of same size, colour, style etc. This is not true of raw feel awareness and electrical discharge in nerve cells, as they are entirely disparate. It is also difficult to accept what is meant when one says that electrical discharge in nerve cells causes this raw feel of awareness, as it is not when one says that the discharge causes a performance or readiness for such performance in the organism. For causal relationship obtains between two events which are quantitatively variable and between which an equation of variation is obtainable. Raw feel of awareness however is nothing but a state of private experience which consists in being consciously aware that *p*, and I can give no sense to the idea of quantitative variation of this *vis-à-vis* electrical discharge.

One of the reasons why neurologists, psychologists and philosophers feel that we shall one day hit upon a brain state which causes or represents consciousness is this. Old-fashioned introspective psychology was based on the assumption that physical phenomena in the brain had psychological causes accessible to introspection. But neurology has revealed that many pieces of human behaviour which introspective psychology thought could only be explained on the assumption of psychical causation, can perfectly be understood in terms of neurological antecedents between which antecedents and performances there is a law-governed relationship. It is then naturally thought by neurophysiology that such a thing as psychical causation is non-existent and its assumption only bars the way to a fuller neurological explanation which, unlike psychical causation, is not based on ghostly entities. Neurological explanation is not only open to interpersonal observation and therefore to the application of scientific rigour, it can actually show the connection that obtains between one state and another by means of laws of physico-chemical functioning. Of course, it is by no means the case that all human behaviour has been thus neurologically explained, but as we now have the correct methodology for study of behaviour – and no false assumptions – we can expect indefinite progress, and as a result a complete neurophysiological science.

Now I do not at all think that a complete neurophysiological science is impossible or that such a science would not be immensely valuable for our understanding of human behaviour. But what precisely has a complete neurological explanation of human behaviour, if obtained,

to offer? It would show that a piece of human behaviour is the result of a certain neural activity between which activity and behaviour is a law-governed relationship. This neural activity itself would be connected in this law-governed manner with other neural activities and so on, so that a complete account of a piece of human behaviour could be stated in neurological terms. But this neurological description could be called complete only as a qualifying description of neurology itself insofar as all possible connections between neural states will have been found. Such a neurology will also show that every piece of human behaviour is a physical manifestation of some neural activity. But it need not be on these scores a complete account of all that can possibly be involved in human behaviour because it leaves out the fact of consciousness and those aspects of human behaviour – these will not be studied because the assumption is that they do not take place – which can take place because there is conscious awareness, particularly behaviour which takes place as a result of choice and decision. It is true that a neurological counterpart of the activity that follows decision and choice can be found, as well as that this neurological activity can be related to other neurological facts of the organism, but there may be instances of human behaviour where all this, even though necessary, will not be a sufficient explanation of behaviour that takes place (or will not be a complete account of what happens).

Let me try to give an example of what I mean. The door of my room is open, I feel discomfort because of draught, get up and close the door. Let us say that here the bodily activity *p* (getting up from the chair) is caused by neural activity *q* (representing feeling of discomfort) which again is caused by neural activity *r* (representing the action of the draught on the body). Here we can talk about a law-like connection between discomfort caused by draught and getting up to shut the door and this will be a complete explanation of my action. Any conscious awareness that I may have of discomfort is not a link in this causal chain – I may not even have any such conscious awareness – it is rather the neural counterpart of 'discomfort' (a psychical term that represents performatory and dispositional tendencies of brain states) which belongs to this chain. But let us now take a more complicated case. There is discomfort through draught and propensity to close the door but at the same time there is expectation that a friend, whom one needs to meet, would shortly be passing by, and a propensity to keep the door open in order to watch the corridor. Here there are present in the organism at the same time two sets of law-like connections between

neural and bodily activities, both in a state of readiness to result in action of a certain sort; $p \to q \to r$ (draught→discomfort→closing the door) and $m \to n \to o$ (need to meet the friend→expectation→keeping the door open). In a case where discomfort and the need to meet the friend are more or less of equal value making it difficult for the agent to decide what to do, the two aforesaid law-like connections are suspended in a state of potential readiness. Both, through this law-like nature, have a tendency to be realised in actual activity yet both cannot be realised at the same time. Suppose then the agent decides to keep the door open for another five minutes, then it is this decision which sets in motion or initiates the actual operation of one set of law-like connections between neural and bodily events by cancelling the other (represented by draught→discomfort→closing the door). Here too the neurologist could give a neurological account of the act of not closing the door (if we can call that an act) since this is connected in a law-like manner with other antecedent events in the organism. And yet this would not be a complete account of what happened, and it may not have happened at all given the same organism and its states, if the decision had been to shut the door and such a decision could be taken considering its connection with certain given antecedent states in the organism. Had the decision been to shut the door the neurophysiologist could still give a neurological account but that would not have been a complete account either, as it would be in a case where I feel no need to keep the door open – a need which I decide to overlook as in this case – and shut it because of discomfort caused by draught. That is because in this case the neurophysiological connections that come into operation have been switched on by a non-neurophysiological event, a decision. Nevertheless, the neurophysiologist need take no account of the decision within his terms of reference and will talk entirely in terms of neurophysiological connections that are available as explanation.

What do I mean when I say that a decision switches on a neurophysiological connection? I can only use an analogy in trying to clarify what I mean. Take the role of the hand which pushes a button and brings on the light in a dark room. The light (in this case) did not switch itself on and it needed a hand to start functioning. Nevertheless, between the button and the light bulb there must already be a connection established through the use of electrical wiring, otherwise nothing would happen when the hand does its job. The electrician knows that the hand is needed to push the button, but if asked how the light

operates he would give his explanations entirely in terms of electricity and wire connections. He would hardly mention the hand and he would be quite right. But for someone who has never seen an electric switch before it may be useful information to be told that one puts the light on by pressing it. This would of course not be a complete explanation of how the light comes on, still, that does not mean that the hand is not playing a part in the total process which is obtaining the illumination that we need. The neurophysiologist, like the electrician, can give a perfectly good explanation of an action, even when that is in some sense a result of decision, entirely in terms of physical processes and their connections, but this would not prove that decision plays no part in the activising of certain processes in the brain, which it switches on, so to say.

A neurophysiologist would no doubt object to this. For him a complete neurophysiological account does not just mean that certain bodily activities are the results of certain neural activities. All neural activities are for him the result entirely of antecedent neural activities and there is no room in this picture of anything like decision, if by that is meant a non-neural something. The neurophysiologist would say that the decision to keep the door open for another five minutes as against shutting it straightaway, must have been the result of some neural activity other than those which were in a state of balance causing indecision and inaction, and it is because of this neural activity that the door was kept open for another five minutes. This seems to me purely speculative, and this speculation is inspired by the dogma that a complete neurophysiological explanation of everything is obtainable. I have discussed the problem of decision more fully in the section on freedom, but let me just say a few words as to why I think that the dogma of a complete neurological explanation of decision is speculative. Human situations of action, particularly those which involve indecision, are different from one another and each such situation insofar as it is different must be governed by a different law. Neurophysiologists must then give a series of laws connecting the relative strength of different degrees of discomfort caused by draught arousing the tendency to close the door, with that of different degrees of urgency and expectation to meet someone arousing the tendency to keep the door open and their possible outcomes. Considering the indefinite variety of situations and persons such laws can hardly be established. They have to be inferred from actual actions and thus would be no more than speculative in nature.

To go back to a point made earlier. A raw feel awareness of experience$_2$, we have said, may be correlated with some neurological state without being identical with it. But it may well be that there is no such correlation. Experience$_2$ happens on the fringe of experience$_1$. I see something red when someone says 'pass me the red book please', and I do so. But I may become consciously aware while handing the book over that it has a red cover. Here I am having a raw feel awareness that something is red, that is, experience$_2$, along with experience$_1$, that is, seeing red. A neurologist looking into my brain when I am having both these experiences will no doubt detect some definite neural activities. But these activities may simply represent experience$_1$ that is, seeing red and not experience$_2$ that is, raw feel awareness that red. Since experience$_2$ happens along with experience$_1$ it is difficult to make sure that neural activities being studied refer to it, let alone determine it causally. What one has to do is to study experience$_1$ (seeing red) in cases where conscious awareness that red is absent and compare the neural states and processes obtained with those involved in experience$_1$ accompanied by experience$_2$. Even if some significant difference appears which definitely connects some neural states and processes with experience$_2$ it is not clear whether at this level there is not just psychophysical parallelism, that is, the presence of both neural and mental states without one causally determining the other. This may be true without there being any parallelism elsewhere. There may be even interaction, but in the way of the mental causing the neural rather than the other way. For experience$_2$ is radically different from experience$_1$ and although it comes in the wake of experience$_1$ that does not prove that it happens in the same way as does experience$_1$ that is, caused by neurological events. Its dependence on neurology may be only this; that its manifestation is parasitic on experience$_1$ which is neurologically determined. Experience$_2$ itself may only be manifested there as a felt state as a result of neurological activities which cause experience$_1$. That is, neurological factors which cause experience$_1$ may only provide suitable conditions for the appearance of experience$_2$ without producing it. If this be the case, and there is no *a priori* reason why it cannot be, then it is possible that there are two different principles of functioning, interacting, but not reducible to one another, namely, the mental where mental means the conscious and the material. This, however, need not in the least upset behaviourist psychology or the hope for a complete neurological science, for behaviourist psychology and neurology study experience$_1$ which, according to me, can very

profitably be identified with neurological events and their behavioural expressions. Even when experience$_2$ has an influence on the doings of human beings such influence does not upset or cause a breakdown of neurophysiological connections, and it may lie solely in a selection out of alternative neurophysiological possibilities.

THE IDENTITY THESIS

Dualists believe that the data of immediate experience are not reducible to or definable in terms of purely physical concepts and the occurrence of experience is not predictable or explainable on the basis of physical laws and descriptions. Although I am a dualist of a sort, according to my analysis immediate experience is experience$_2$ and it consists solely in a state of conscious awareness that *p*. If it has any data, that is nothing other than the data of experience$_1$ on the fringe of which it occurs. This experience$_2$ is not definable in terms of physical concepts but dualists often claim that even experience$_1$ consists of more than neurological events and their behavioural manifestations, and that they have phenomenological or experienced counterparts. I agree with behaviourists that events belonging to the order called experience$_1$ do not have phenomenological counterparts of a psychical nature, even though I want to insist that there can be such a thing as immediate experience or experience$_2$ of behaviourally studied events, which has phenomenologically a distinct character of its own. However, this experience$_2$ has no features of its own apart from features of experience$_1$ and its only characteristic whenever it occurs is that it is a state of conscious awareness that *p*. The identity thesis makes no distinction between mental$_1$ and mental$_2$ and asserts that the mental is nothing other than the physical. Let us now look at some arguments which various identity theorists have put forward.

Herbert Feigl

What is at issue, says Feigl[1], in the identity thesis is the predictability of mental events from neurophysiological events. It seems that Feigl thinks that there are mental events of a psychical character corresponding to neurophysiological events, only he wants to say that they are identical with neuralogical events in some sense the exact nature of which is not quite clear to me. Of course mental events in the sense of

[1] 'The "Mental" and the "Physical"' in *Minnesota Studies in the Philosophy of Science*, Vol. II.

$mental_1$ are predictable from neurological events and as $mental_2$ appears when the brain is in a state of high excitability resulting from $mental_1$ it can be expected to be there on such occasions, but that would not make it predictable from neurological events if by that is meant that it is produced by the latter. If Feigl's mental events are states of conscious awareness that such and such, which occur when mental events occur, then they would not be predictable in this sense. But it is not quite clear what Feigl's mental states stand for.

Feigl makes a distinction between the concepts of $physical_1$, which refers to inorganic processes, and $physical_2$ which uses concepts, methods and laws appropriate to the study of intersubjective or public phenomena. Feigl starts with saying that the mental events are physical events in the sense that they can be studied by intersubjective methods but ends by identifying them with physical events, that is, neural events in the brain. Anyway, his objection against interactionism is that it involves a breach of physical determinism, for if mental events directly cause neural events then these neural events cannot be seen as correlated with other neural events, and a breach in determinism – a concept that belongs to $physical_2$ language – takes place as a result of which a complete science of neurophysiology becomes impossible. This would no doubt be so if mental events directly cause neural events or create neural events out of the blue, not, as I have argued, if they only activate or switch on some already existing neurological connections in preference to other connections which are also there in the organism in a state of readiness. The neurological organism being immensely complex in structure and function and being subject to diverse stimuli at the same time, may hold more than one possibility of action at any one moment. A $mental_2$ event then may influence the actual occurrence of a neural event without causing it in the full sense, that is, a mental event may be necessary for the occurrence of a particular physical event that takes place but not sufficient.

Even this may be thought to be a breach of physical determinism insofar as physical determinism does not show such $mental_2$ influence elsewhere than in the human organism. But then physical phenomena cannot show this influence as they, unlike human beings, do not exhibit mental characteristics. Feigl however admits that a human personality exhibits mental characteristics which are presumably of $mental_2$ order since he talks about raw feel. This is why he takes his stand on $physical_2$ determinism which he says is a methodological requirement that things can only be said to be sufficiently explained

when they are shown to be correlated with other things open to an intersubjective process of observation. This I believe can be done in respect of all bodily phenomena whether or not they can be said to be influenced by mental phenomena, for even when they are so influenced the neurological processes involved are such that they are made possible by the presence in the organism of other neurological states and processes with which they are correlated. Unless this were so the mental phenomenon would have nothing to exert its influence on, no foundation on which it could work, a foundation that has to be laid in the organism by way of definite neurological connections. Since such neurological connections are there in respect of all bodily activities whether or not they are influenced by choice, total breach with physical determinism need not be feared. The only breach is in respect of $physical_1$ language as advocated by old-fashioned materialism which says that there is nothing at work in the world but laws of physics and chemistry. But it may well be that physical language is not applicable to consciousness, and there is no compelling reason why something to which $physical_1$ language does not apply cannot be there.

Scientists, says Feigl, are opposed to the admission of purely subjective factors as a basis for prediction. This is understandable, and as most human behaviour is what it is because it is conditioned in certain ways and as human behaviour does exhibit law-like patterns, a good deal of human behaviour remains in principle predictable. The only instances of human behaviour which are not in principle predictable are those which involve a genuine need for choice and decision, that is, where one has to make up one's mind as to which of the available alternatives to accept as one's goal before one acts. Such instances do not constitute more than a small segment of human behaviour but they remain intersubjectively unobservable and thus beyond the scope of scientific methodology.

Unlike Farrell, Feigl does not deny that there is such a thing as raw feel experience which can be an object of direct introspection but he says that it is also $physical_2$ in the sense of being intersubjective in as much as people can still talk about these experiences and communicate them to others. Feigl's raw feel is unlike what I have called raw feel awareness which consists in experience that *p*. His raw feel looks like a psychical counterpart which consists in experience of red rather than experience that red. So conceived it certainly can be studied by $physical_2$ language for when one describes this so-called $experience_2$

of red, one is really describing experience of red which can be intersubjectively studied. Raw feel then is physical in the sense of $physical_2$ which is a matter of methodology and intersubjective approach to facts. But it does not follow from this, as Feigl seems to think it does, that raw feel experiences are also physical in the sense of $physical_1$. This is done when he asserts that these are 'identical with neurological events'. He claims that his theory is not old-fashioned materialism, and also that a satisfactory solution of the relation between mental and physical must render an adequate account of the efficacy of mental states but I do not see that these claims come to anything when mental states are finally identified with neural states. Of course he comes to say that mental states are $physical_1$ in nature via saying that they are physical in the sense of $physical_2$ but I am unable to see that it makes any real difference.

So I confess that I find it difficult to figure out in what the identity thesis as advocated by Feigl consists. It is supposed to be different from a straightforward materialistic thesis that consciousness is nothing but a neurological event in the brain, presumably because Feigl admits to raw feel experience. But even though Feigl does not use the reductionist formula 'nothing but' he still says that raw feel experience is identical with a neurological event. But as raw feel experience has an existential identity of its own I do not understand what he means by saying that it is identical with a brain state. Surely this does not mean numerical identity – for if it means 'caused' then we shall have to say that causes and effects are identical, with the uncomfortable result that we shall have to say that according to scientific determinism the whole world is an identical unit. Anyway, the identity in question is not logical identity, says Feigl, that is, statements about mental states are not statements about neural states and these two kinds of statements are not logically translatable into each other. If mental and neural states were logically equivalent there would be a necessary relationship between mental state ψ and neural state φ and it would not have been empirically necessary to find out which mental state corresponds to which neural state. The relationship is logically contingent and it just happens to be the case that *p*, a mental state, corresponds to *q*, a neural state.

Now it seems to me that Feigl treats the ψ φ relationship as contingent only so far as one has to find out by empirical search which mental state corresponds to which central state, that is, whether *p* corresponds to *q* or to *r*, *s*, *t*, etc. But the existence of this sort of identity between ψ and φ seems to have been logically taken for granted and hence in some sense treated as necessary in the very claim that $physical_2$ determinism and consequently $physical_1$ language are all that are necessary for the understanding of raw feel experience. Unless such is the case the programme of studying a raw feel state by studying a neural state cannot be justified. In this programme any doubt as to whether a conscious state can be identified with a neural state cannot be entertained. So this point about logical contingency as regards which mental state corresponds to which neural state does not seem to make much difference in distinguishing Feigl's thesis from a reductionist materialism which presumably also admits that empirical search is necessary for finding exact identifications. For despite this contention about logical intranslatability Feigl concludes that mental states are simply central states. Logical intranslatability means that to talk about a feeling is not to talk about a brain state even if they are identical. But it seems to me that if to talk about *x* is not to talk about *y* then there is an obvious sense in which *x* and *y* are not identical. Moreover the logical gap between the two languages should not be insurmountable if only one type of entity is involved as referents. At the moment we do not happen to know the central states with which our phenomenal states are identical. There may however come a time, with increase in neurophysiological knowledge, when a precise and complete list of correspondence between mental and neural state would be available for universal consumption and when such knowledge would become a part of our commonsense understanding of ourselves. The logical gap should then be overcome and we should be able to use these languages interchangeably. If we cannot do so despite this knowledge, then these languages must be taken to have genuinely different referents belonging to different orders of experience and this is what creates a logical gap between them. According to the account that I have given of $experience_2$ which consists of a state of conscious awareness that *p*, knowledge of $mental_1$ and neural identities would not close the logical gap. For when I talk about $experience_2$ what I *intend* is the raw feel of awareness involved in conscious experience and I do not intend any neural event involved in the corresponding $experience_1$ of which I may be perfectly aware and which I may even be observing on a cerebroscope.

If we had, says Feigl, completely adequate and detailed knowledge of the neural processes in the human brain and the knowledge of the one-one or at least one-many ψ φ correlation laws, then the description

of a neural state would be a completely reliable guidance or a genuine criterion for the occurrence of a corresponding mental state. This seems to me perfectly true. If by mental is meant $mental_1$, and it is only the $mental_1$ which can be studied scientifically or otherwise, then knowledge of correspondence between $mental_1$ and neural is surely a sufficient justification of our inference regarding the mental on the strength of the neural. Not only this, as $mental_2$ or $experience_2$ occurs on the occasion of $mental_1$, the presence of $mental_1$ supported by verbal claims is sufficient justification for us to believe that one has a mind capable of both $experience_1$ and $experience_2$. Of course $experience_2$ remains private in the sense that raw feel of conscious awareness is such that it can be felt only by a specific individual who is having the experience. But this is no reason why understanding of its presence in any centre other than oneself cannot be had when that centre communicates such presence by verbal means. Verbal communication regarding $experience_2$ at least makes that experience publicly referable without however making it public in every sense of the term – (meaning that the description 'private' still remains applicable to it despite this public referability).

U. T. Place

Consciousness, says Place[1], is either a special type of behaviour, running back and forth behaviour, or a disposition to behave in a certain way. Concepts like knowing, believing, wanting, intending can be analysed in terms of dispositions to behave. Concepts like sensations, images, consciousness, experience seem to require inner processes or events. Even if we can talk about inner processes, such processes do not form a separate category over and above the physical and physiological processes with which they are known to be correlated. However, Place agrees with Feigl that although consciousness is a brain process statements about sensation and mental images are not reducible or analysable into statements about brain processes in the way cognition statements are analysable into statements about behaviour.

According to Place objection to the identity thesis arises out of a failure to distinguish between the 'is' of definition and the 'is' of composition. In the case of 'red is a colour' (definition) there is a relationship between the meaning of the expression forming the grammatical predicate and the meaning of the expression forming the

[1] 'Is Consciousness a Brain Process?' *The Philosophy of Mind*, ed. V. C. Chappell (Prentice-Hall, Englewood Cliffs, N.J., 1962).

grammatical subject such that whenever the subject expression is applicable the predicate must also be applicable. In the case of 'His table is an old packing case' (composition) it merely so happens that in this case both expressions are applicable to and at the same time provide an adequate characterisation of the same subject. According to Place,

> those who contend that the statement 'consciousness is a brain process' is logically untenable base their claim, I suspect, on the mistaken assumption that if the meanings of two statements or expressions are quite unconnected, they cannot both provide an adequate characterisation of the same object or state of affairs: if something is a state of consciousness, it cannot be a brain process, since there is nothing self-contradictory in supposing that someone feels a pain when there is nothing happening inside his skull.[1]

It seems to me that Place is not being fair to the opponent of the identity thesis. The objection surely is not that *if* two descriptions are logically quite independent of each other, as those of being a table and being a packing case, they cannot characterise the same state of affairs. The 'same state of affairs' surely means here that there is an existentially identical (and identifiable as such) something, *x*, which can be said to be both a packing case and a table, although being a table and being a packing case are independent states of affairs. Take for instance the statement, 'the brother of Y is a friend of Z'. Being a brother is not logically derivable from being a friend, but nobody would find the statement objectionable since there can be an identical individual, X, who is both a brother of Y and a friend of Z. 'Consciousness is a brain process', if by consciousness we take a state of conscious awareness experienced as such, is unlike 'A brother of Y is a friend of Z' or 'his table is an old packing case' in that it is not obvious that what is involved here is an existentially identical state of affairs which is being characterised both as consciousness and as brain process. If we take consciousness as intelligent performance, as is often the case with behaviourist approach to the problem, then of course it is perfectly in order to say that what is involved is an existentially identical process which is neurophysiological (a neural state culminating in a certain type of bodily activity or response). But if we take consciousness to be $experience_2$ then there are clearly two existentially differentiable processes involved, raw feel experience and neural process, and

[1] Ibid, 103

whatever be the relationship between the two it is not existential identity as in the case of 'a brother of Y is a friend of Z'.

An intersubjective scientific process, says the materialist, cannot allow anything which is not knowable except through introspection. What exactly does this mean? Raw feel experience cannot obviously be an object of study of an intersubjective scientific method. Not only can such a method not be applied to it, there is nothing here that can be studied in detail since there are no details to study. Raw feel experience is a function of awareness that *p* as well as of itself and this is all there is to it. But the fact that consciousness cannot be studied by a scientific method does not mean that science cannot allow that consciousness exists at all as a distinctive sort of function. Besides, the privacy of raw feel awareness does not mean that it is not publicly referable and communicable in interpersonal talk; how else do we manage to carry on an intersubjective discourse about it? However, as there are no states of consciousness, all states being mental states of which we can become aware in being conscious of them – this may be referred to as 'a state of being conscious' and it is the same *qua* such a state no matter what one is conscious of – there is no question of its being an object of scientific study in the same way that the neurophysiological organism and its different states and processes can be.

To get back to Place. Let us take Place's example 'His table is an old packing case' and see in what sense it is a statement of identity. Obviously we cannot say 'A table is a packing case'. We cannot do this because as usage stands now the referent of 'packing case' is not existentially identical with the referent of 'table'. 'His table is an old packing case' means presumably that he is using an old packing case as a table and insofar as this is the case we can call it a table. It does not matter whether we call the object a 'packing case' or a 'table', for there is an existentially identical object which can be put to different uses and accordingly named differently. In the case of 'consciousness is a brain process', it is the very idea of there being an existentially identical thing for which both the names 'consciousness' and 'brain-process' are suitable descriptions, which is in dispute. Non-behaviourists find that there is clear evidence of there being two utterly dissimilar things or events when they are confronted with consciousness and brain process and this is why they object to the identity thesis and not because they confuse the 'is' of composition with the 'is' of definition.

The argument, says Place, from the logical independence of two expressions to the ontological independence of two entities breaks down

in the case of consciousness and brain process just the same as in the case of 'A cloud is a mass of tiny particles usually water droplets in continuous motion'. A cloud remains 'a fleecy texture suspended in the atmosphere whose shape is subject to continuous and kaleidoscopic change'. This is certainly not logically the same as 'a mass of tiny particles usually water droplets in continuous motion'. Nevertheless a cloud is not a different entity from a mass of particles etc. But the argument in the case of our objection to 'consciousness is a brain process' need not be from the logical independence of expressions to ontological independence of entities. It can be the other way, that is, from the ontological independence of entities – completely disparate nature of two different orders of entities – to the logical independence of expressions. The argument goes like this: 'consciousness' and 'brain process' refer to two different kinds of states of affairs because there are two different kinds of states of affairs to symbolise which these two different expressions have been coined – and not like this; as 'consciousness' and 'brain process' are two different expressions there must be two different things to which they refer.

Place supports his thesis of identity of consciousness and brain process by exposing what he calls 'the phenomenological fallacy'. This is the mistake of supposing that when the subject describes his experience, when he describes how things look, sound, smell, taste and feel to him, he is describing the literal properties of objects and events on a particular sort of internal cinema or television screen, usually referred to in psychological literature as 'the phenomenal field'. I agree with Place that such a belief is a fallacy. But phenomenologists who believe in the reality of consciousness as something *sui generis* do not generally adhere to the doctrine of the phenomenal field, and the display of things on an internal screen. Sartre, for instance, rejects the idea of there being anything *in* consciousness. Consciousness is a function of awareness and revelation whose sole reality consists in intending an object which is given out there, as well as in being aware of itself as such a function. If there is a phenomenal field it is out there consisting of objects intended by consciousness, it is never inside consciousness. But of this more later. All I want to point out here is that the absence of an internal phenomenal field by no means proves that consciousness is nothing but a brain process. Place however says that once we rid ourselves of the phenomenological fallacy 'we realise that there is nothing that the introspective subject says about his conscious experiences which is inconsistent with anything the physiologist

might want to say about the brain process which causes him to describe the environment and the consciousness of that environment in the way he does'. When the subject describes his experience by saying that a light which is in fact stationary appears to move, all that the physiologist or the physiological psychologist has to do in order to explain the subject's introspective observations is to show that the brain process which is causing the subject to describe his experience in this way is the sort of process which normally occurs when he is observing an actual moving object and which therefore normally causes him to report the movement of an object in his environment.

This clearly shows that all that Place means by conscious experience is experience which is the actual response that an intelligent subject makes verbally or otherwise to a stimulus. It becomes the datum of introspective observation when the subject attends to his own response and talks about it. If all that Place has in mind in talking about consciousness is this verbal response then as the ability to respond verbally can be identified with certain brain processes consciousness can be so identified too. But this leaves out raw feel awareness of the subject when there is this awareness that he is attending to his own response and describing it. It is not so clear that this can be identified with the brain processes that are involved in verbal response.

J. J. C. Smart

Let us now examine the arguments advanced by Smart in 'Sensations and Brain Processes'.[1] According to Smart science is increasingly giving us a viewpoint whereby organisms are able to be seen as physico-chemical mechanisms and it seems to him that all behaviour of man will eventually be explained in purely mechanistic terms. Now I want to point out the difference between saying that (i) the human organism is a physico-chemical mechanism and that (ii) a human being is nothing but a physico-chemical mechanism. An interactionist who says that human consciousness influences some activities of the human organism does not have to deny that the brain and the nervous system constitute a physico-chemical mechanism. All he has to say is that a human personality is a psycho-physical system, as distinct from the human organism which is a physico-chemical mechanism, and in this system consciousness sometimes plays a part in switching on, as it were, some electrical events, without however creating them, when the events switched on are related to other antecedent events in the

[1] In *The Philosophy of Mind*, ed. V. C. Chappell.

organism. The fact that a human personality is a psycho-physical system, if it be a fact, does not then require us to reject that the human organism is a physico-chemical mechanism. But of course it does require us to believe that this mechanism, unlike other mechanisms of physico-chemical nature, can be acted upon by something which is not physico-chemical. However, this need not contravene well-established physico-chemical laws or neurophysiological connections if the complexity of the neurophysiological mechanism and its openness to variety of stimuli at the same time allow it to hold alternative possibilities between which a choice may be made.

What Smart wants to say is that the human personality is nothing but a physico-chemical mechanism. According to him, 'That everything should be explainable in terms of physics (together of course with descriptions of the ways in which parts are put together – roughly, biology is to physics as radio-engineering is to electromagnetism) except the occurrence of sensations seems to me to be frankly unbelievable'.[1] Sensations here of course mean sensations$_2$ and not just responding to stimuli. It is not clear why this is unbelievable – there are people who believe exactly this and think that they do so with good reason – except in the light of the assumption that everything must be explainable in terms of physics. But the proposition 'everything must be explainable in terms of physics' is not logically necessary. Yet Smart inclines to take it in this manner, or at least as an empirically well-established theory, which it is not.

The thesis that sensations are brain processes, says Smart, is not the thesis that 'after image' or 'ache' means the same as 'brain process of the sort X'. The thesis is

> . . . that insofar as 'after image' or 'ache' is a report of a process, it is a report of a process which *happens to be* a brain process. It follows that the thesis does not claim that sensation statements can be *translated* into statements about brain processes. . . . All it claims is that insofar as *a sensation statement is a report of something* that something is in fact a brain process.[2]

Here Smart makes two remarks: (1) Insofar as it is a statement about a process the process is a brain process and (2) Insofar as it is a statement about *something* that something is a brain process. I think that the first remark is acceptable while the second is not. 'Sensation' when used as a

[1] Ibid., 161.
[2] Ibid., 163.

mental concept is supposed to include the raw feel of experience$_2$. Experience$_2$ is not a process, it is a state of conscious awareness that something is the case, such as pain in the stomach, when pain can very well be the result of a brain process. There is no process involved in the conscious awareness of pain, *qua* awareness, although there is a process involved, a brain process, in the occurrence of the sensation of pain of which sensation$_2$ is an awareness. So insofar as a statement about a sensation is a statement about a process, that process is a brain process, but a statement about sensation may be about the experiencing of pain, that is, the experience of having it with emphasis on having rather than pain as a process of a distinct sort. Then the statement is about something which is not a brain process.

So the statements about sensations which are not translatable into statements about brain processes are statements about sensation$_2$. Sensation$_1$ statements seem to me perfectly translatable into brain process statements where a sensation$_1$ statement is a statement about a discriminative performance such as is involved in responding correctly to a signal which is red, or uttering the word 'pain' when in a certain state of physical disturbance. The seeing of red and the feeling of pain – these are mental$_1$ terms – such as are involved in these responses, happen because of certain brain processes and if we know what these processes are there should be no difficulty in translation. The difficulty is with regard to mental$_2$ terms where seeing or feeling is also a state of conscious awareness that red or pain.

Let us examine some of Smart's examples where a logical gap between statements does not signify different referents. One of his examples is statements about nations and individuals. Their logic is different but there is no such thing as a nation over and above individuals. But insofar as a statement about a nation cannot be translated into a statement about individuals the two statements do have different referents. Admittedly there is no entity called a nation comparable to an entity called an individual but the concept 'nation' does stand for something over and above individuals each taken singly. Individuals, as they come under the concept 'nation', are thought of as forming a collectivity or a unit for a political or some such purpose. If these different referents were not involved the logic of the two statements would not have been different in the way they are. Take again the statement 'the morning star is the evening star'. If the logic of the morning star is different from the logic of the evening star despite this identity statement, as Smart admits that they are, that is because two

distinguishable referents are involved, *the star as seen in the morning* and the *star as seen in the evening*; only it happens to be contingently the case that they are existentially the same entity. So there is a sense in which it may be said that the morning star is not the evening star, that is, if we want to emphasise the time when we see the star. If the star as seen in the morning is named '*x*' and the same star as seen in the evening '*y*' then it would obviously not do to say that *y* is seen in the morning. So despite this existential identity a statement of identity about the morning and the evening star may be false if we have two logically distinguishable referents in mind. But in the case of consciousness and brain process we do not even have this existential identity.

Take again the statement 'lightning is electrical discharge due to ionization of clouds of water vapour in the atmosphere'. 'Lightning', says Smart, is not translatable into 'electric discharge etc.' It may not be so translatable for an uninformed person who knows nothing about electrical discharge. But for someone who knows what 'electrical discharge due to etc.' means and who knows that lightning is produced by this, the two concepts should be translatable. But lots of people who know a great deal about neurophysiology – some reputed neurophysiologists, for instance – find it impossible to identify consciousness (raw feel experience) with a brain process, because there obviously appears to be a categorical distinction involved, a distinction between two orders of reality, one physical and the other non-physical, not just two different things of the same order, like lightning and electrical discharge, one of which, lightning, can be a manifestation of the other, electrical discharge, in certain specified conditions. I am not arguing that sensations cannot be brain processes on the ground that someone who knows nothing about brain processes can still talk about sensations. But if a man who is in a state of conscious awareness that red (sensation$_2$) watches the neural event representing perception of red (sensation$_1$) on a cerebroscope and finds that the two states involved are of a different order, one spatial, the other non-spatial etc., then it becomes difficult to understand what exactly is meant when one says that they are identical. Such difficulties would not arise in the case of cloud and lightning for people who know enough physics to understand how they are produced, and for them a cloud statement would be translatable into a water droplet statement, or a lightning statement into electrical discharge statement without much trouble.

Smart recognises, unlike Place, that whether consciousness is identical with a brain process or not is not an empirical issue and we

cannot therefore pin our hope of establishing this on future increase of knowledge. For no matter what neurophysiology has to tell us it will always be possible to insist that consciousness is not identical with a neural process. This will be possible, so it seems to me, because of the nature of consciousness, which unlike an electrical discharge or a performance of the organism, is present to itself – consciousness that *p* is also consciousness of itself – and is immediately felt as such. This makes the identification of consciousness with electrical discharge quite unlike the identification of morning star with evening star or of cloud with water droplets. It can hardly be denied that there is consciousness and that consciousness does appear to be something quite different from electrical discharge; if it did appear to be the same as electrical discharge – as lightning does after explanation even if it did not do so initially – there would have been no need for so much argument by identity theorists to prove their identity. Perhaps an identity theorist would say that the appearance of consciousness as something quite different from electrical discharge is a grand illusion from which we suffer and which can be got rid of by means of right knowledge – not empirical knowledge (which Smart is aware cannot settle the issue) but shall we say metaphysical knowledge – and that there is ultimately nothing but physico-chemical processes in the universe whatever else may also *appear* to be there. This will be the Indian doctrine of illusion in reverse which says that there is only consciousness in the world and the appearance of this consciousness as the physical world is a grand illusion which can be got rid of by means of right knowledge (this also is metaphysical in nature).

Richard Rorty

Richard Rorty in his article 'Mind-Body Identity, Privacy and Categories'[1] says that there are two forms of the identity thesis, the translation form and the disappearance form. To say that a sensation is identical with a brain process involves this difficulty: a sensation of pain may be stabbing but the brain process cannot be described in this way. So philosophers have criticised the identity thesis by saying that it involves a category mistake. Rorty contends that the theory can be stated in such a way that it does not involve category mistake. One may try to do this with the translation form. A statement like 'I am having a stabbing pain' could be translated into topic-neutral language of the form, 'there is something going on which is like what is going on when

[1] In *Philosophy of Mind*, ed. Stuart Hampshire (Harper & Row, New York, 1966).

etc.', and then it could be said that there would be no difficulty in identifying it with a brain process. But it is not easy to find topic-neutral language for sensations, so Rorty suggests that we adopt the disappearance form in explaining what identity means in this context. According to Rorty it is the sort of relation that holds e.g. between quantity of caloric fluid and mean kinetic energy of molecules.

> There is an obvious sense of 'same' in which what used to be called 'a quantity of caloric fluid' is *the same thing* as what is now called a *certain* mean kinetic energy of molecules, but there is no reason to think that all features truly predicted of the one may be sensibly predicted of the other.[1]

As talk about caloric fluid has disappeared with the progress of science leaving only talk about kinetic energy of molecules, so with the future progress of psychophysiology talk about sensations of pain could disappear leaving only talk about brain process. Rorty does not think that this is going to happen because of the inconvenience involved but plainly he rejects the thesis of the logical intranslatability of the language of sensation into language of brain process advanced by Feigl, Place, and Smart.

Sensation statements no doubt have a reporting function which brain statements do not have. But according to Rorty what one is really reporting when one is reporting pain is, 'my c-fibres are firing', and this will become clear when we have acquired enough knowledge of neurophysiology. Both the explanatory and reporting functions of statements about sensations can be taken over by statements about brain processes. To see this is to eliminate useless questions about sensations from neurophysiological enquiry, leaving however our ability to describe and predict undiminished. This seems to me perfectly valid, for it is much more profitable for neurophysiology to ask questions about brain processes in order to see their interconnections as well as their role in our actual responses than to ask questions about mental items. But it is at the same time true that neurophysiological questions are not the only questions to ask about human behaviour and there may be a different framework of enquiry in which questions about sensations$_2$ or consciousness may become meaningful and this may have as much relevance for a philosophy of mind as neurophysiology. Anyway according to Rorty we could convey all we convey by

[1] Ibid., 33.

'I am in pain' by saying instead 'my c-fibres are firing'. Not only that; 'my c-fibres are firing' is the information that I am really conveying when I say 'I am in pain' although I may not know this to be the case.

Some philosophers have objected to the suggestion that to say 'I am in pain', is the same thing as to say 'my c-fibres are firing', and even some identity theorists have found a logical gap between them. But it seems to me that when it does become common knowledge that pain is the result of firing of c-fibres one can very plausibly convey the fact that one is in pain by saying that one's c-fibres are firing. Also, when we do come to have this information at the commonsense level there should be no difficulty in seeing that at least one of the things I am reporting when I am saying 'I am in pain' is that my c-fibres are firing. I do not have to have a periscope hitched up to a microscope to give me a view of my skull in order that I may report a brain process, once I know through a thorough knowledge of neurophysiology which brain process stands for which sensation. Nevertheless it seems odd to say that a man who has never heard of brain processes and who is saying 'I am in pain' is really reporting here and now about his brain processes, even though it is perfectly correct to say that he would not be reporting what he is reporting unless he had the appropriate brain process, as well as that he could be made to see, by supplying him the necessary information, that there is a sense in which what he is reporting is the firing of his c-fibres. Thus it is perfectly possible to report in a sensation statement public events such as brain processes, provided both the reporter and the reported understand this language. But it does not seem to me true that one cannot possibly report anything but public entities when one reports sensations.

What then may we say a man is reporting when he is saying 'I am in pain'? He may be reporting that it is pain and not something else that he is having, such as an itch or a ticklish sensation. But he may also be reporting his *having* of the pain, meaning his conscious awareness of the experience of pain that he is having which is distinguishable from responding to pain sensations by exclamations, grimaces etc. If the man's intention is to talk about pain when he says 'I am in pain', his emphasis is on the public reference of the term 'pain' whether or not it is thought to be brain process and insofar as this is the case it is not his intention to report something private to himself. Nevertheless he may be reporting something which is in some sense private – his conscious undergoing of an experience of pain. I do not mean by using the term

'private' that we do not understand what he is talking about when he reports something which is private, it is only private in the sense that his undergoing of this experience cannot be jointly shared by another man although the latter can equally undergo his own experience of pain. It is only private in contrast to a process in his brain which can be intersubjectively looked at. If the man is doing this, if his object is to talk about his experiencing, about his being aware of pain, then he is not talking about a public object as he is doing when he is talking about pain, a mental$_1$ event which consists of firing of c-fibres and a performatory response to it. So a mental$_2$ event, the having of an experience of pain, does not disappear simply because one knows that the occurrence of pain (a mental$_1$ event) means the firing of c-fibres.

I shall in the light of the above consider Rorty's criticism of Baier's[1] 'privacy' objections to the identity thesis. According to Smart we accept Smith's introspective report about his experience in preference to what others say because at the present stage of lack of development of brain process theory we have not yet got a public criterion. When that criterion is developed we shall prefer public reports about experience to private introspection. Baier thinks that this is simply a confusion. Introspective reports are necessarily about something private and this is incompatible with having a public criterion. Baier agrees that physiological theory may so progress that we shall come to have absolutely reliable encephalograph-based claims about people's experiences. We may then know for certain that Jones is having a pain even though Jones says 'no' (he may be lying or concealing the fact of his pain for some reason). But this will never show to the sufferer that he is mistaken in thinking that he is having a pain, for such a mistake is inconceivable. Physiology can thus give us evidence but not criterion for saying that Jones is having an experience of a certain sort. For this, says Baier, the sufferer's epistemological authority is final. Hence introspection talk and brain process talk are not the same.

Rorty points out that if a physiological law connecting a brain process with sensation of pain does become well established (that is, encephalograph reports do always agree with people's reports of sensation of pain) then we are entitled to think that Jones is mistaken if the encephalograph says that he is having pain and he says 'no'. Jones, we would think, must be misjudging or misnaming; he does not know that what he is feeling is called 'pain'. According to Baier 'however good the evidence may be such a physiological theory can never be

[1] Kurt Baier, 'Smart on Sensations', in *Australasian Journal of Philosophy*, 40 (1962).

used to show to the sufferer that he was mistaken in thinking that he had a pain, for such a mistake is inconceivable'. Rorty replies

> We now see that the claim that 'such a mistake is inconceivable' is an ellipsis for the claim that a mistake, made by *one who knows what pain is*, is inconceivable, for only this expanded form will entail that when Jones and the encephalograph disagree, Jones is always right. But when formulated in this way our infallibility about our pains can be seen to be empty.[1]

Now I do not know exactly in what respect Baier thinks that mistake about experience is inconceivable. Is it inconceivable in respect of the particular experience reported (in this case pain) or in respect of the having of the experience? Baier *does not say* that if the encephalograph says 'yes' and Jones says 'no', Jones' 'no' must be taken as a proof that the encephalograph is mistaken. On the contrary he suggests that Jones may be lying. What he says is that a physiological theory can never be used to show to the sufferer *that he is mistaken in thinking that he is having a pain*, for such a mistake is inconceivable. It may be that the mistake that is inconceivable for the sufferer to make is not so much about pain but about *thinking that he is having a pain*. That is, it is inconceivable that one would think that one is having an experience x, with emphasis on having rather than x, when one is not having any experience at all. That is to say, if somebody constructs an instrument designed to detect whether one is having an experience, that is, is in a state of conscious awareness (that x) and this instrument said 'no' when Jones said that he was in a state of conscious awareness (that x) we would have to believe Jones rather than the instrument. But it is not inconceivable that one may make a mistake as regards what the experience is about. It may be in fact an experience y when one says that it is x. It is highly unlikely that such a mistake will be made about pain, but if the experience is a particular kind of pain then one having the pain may not be able to correctly identify as to what kind of pain it is. So we may agree with Wittgenstein's point that sensation reports must conform to public criteria or else be disallowed and yet claim that there is something private about the having of a sensation and that it is impossible for one to believe that one is having a sensation when one is not, although one may be mistaken as to what the sensation is sensation of.

[1] Ibid., 53.

The criteria of privacy that Baier gives seem to me to fit better the *having* of a sensation of pain, that is, conscious awareness that p rather than the sensation itself, that is, pain understood as a certain kind of physical process. For pain understood as a physical process is not imperceptible by the senses if we include the cutaneous sense, but imperceptibility by the senses is one of Baier's criteria for privacy. The others are: 1) necessarily owned, 2) necessarily exclusive and unshareable, 3) necessarily asymmetrical – it makes no sense to say I could see or hear that I had a pain but it makes good sense to say that I could see or hear that X had a pain, 4) necessarily of such a nature that the person having it needs no evidence for it, while others may, 5) it is something about which the person whose private state it is has final epistemological authority. It makes no sense to say 'I have a pain unless I am mistaken'.

If we think of the experience as not just pain but a particular sort of pain – stabbing pain, shooting pain, a throbbing pain, a dull ache or just local sensitivity – it is easy to see that the agent having the pain is not infallible or even necessarily the best judge as to what sort of pain it is, and a sufficiently refined encephalograph report may give us more correct information than the agent's introspective report can. I myself sometimes find it difficult to give an accurate description of a pain when asked by a doctor to do so, but the doctor with his long experience and with the help of an examination of the source of pain may know better how to identify the nature of the pain. Anyway, this shows that the sufferer's epistemological authority is necessarily final not with regard to the accurate nature of the particular experience had but with regard to the having of the experience. Rorty can no doubt say to this that if we are studying human behaviour then the mere information that one is having an experience, even if this is infallible, is hardly of any value. We want to know what sort of experience one is having and for this publicly inspectable states or properties, whenever obtainable, are of greater value than introspective report. This seems to me quite true. But these public properties will only explain the nature of an experience, what it is about, rather than the fact of the occurrence of a conscious having of an experience for which the agent's epistemological authority does seem final. Neurophysiology may have no use for this latter, for its interest lies in finding interconnections between different states and processes, neurological and physiological. But what is useless for neurophysiology may be useful for certain other purposes relevant to what is called 'philosophy of mind'.

might want to say about the brain process which causes him to describe the environment and the consciousness of that environment in the way he does'. When the subject describes his experience by saying that a light which is in fact stationary appears to move, all that the physiologist or the physiological psychologist has to do in order to explain the subject's introspective observations is to show that the brain process which is causing the subject to describe his experience in this way is the sort of process which normally occurs when he is observing an actual moving object and which therefore normally causes him to report the movement of an object in his environment.

This clearly shows that all that Place means by conscious experience is experience which is the actual response that an intelligent subject makes verbally or otherwise to a stimulus. It becomes the datum of introspective observation when the subject attends to his own response and talks about it. If all that Place has in mind in talking about consciousness is this verbal response then as the ability to respond verbally can be identified with certain brain processes consciousness can be so identified too. But this leaves out raw feel awareness of the subject when there is this awareness that he is attending to his own response and describing it. It is not so clear that this can be identified with the brain processes that are involved in verbal response.

J. J. C. Smart

Let us now examine the arguments advanced by Smart in 'Sensations and Brain Processes'.[1] According to Smart science is increasingly giving us a viewpoint whereby organisms are able to be seen as physico-chemical mechanisms and it seems to him that all behaviour of man will eventually be explained in purely mechanistic terms. Now I want to point out the difference between saying that (i) the human organism is a physico-chemical mechanism and that (ii) a human being is nothing but a physico-chemical mechanism. An interactionist who says that human consciousness influences some activities of the human organism does not have to deny that the brain and the nervous system constitute a physico-chemical mechanism. All he has to say is that a human personality is a psycho-physical system, as distinct from the human organism which is a physico-chemical mechanism, and in this system consciousness sometimes plays a part in switching on, as it were, some electrical events, without however creating them, when the events switched on are related to other antecedent events in the

[1] In *The Philosophy of Mind*, ed. V. C. Chappell.

organism. The fact that a human personality is a psycho-physical system, if it be a fact, does not then require us to reject that the human organism is a physico-chemical mechanism. But of course it does require us to believe that this mechanism, unlike other mechanisms of physico-chemical nature, can be acted upon by something which is not physico-chemical. However, this need not contravene well-established physico-chemical laws or neurophysiological connections if the complexity of the neurophysiological mechanism and its openness to variety of stimuli at the same time allow it to hold alternative possibilities between which a choice may be made.

What Smart wants to say is that the human personality is nothing but a physico-chemical mechanism. According to him, 'That everything should be explainable in terms of physics (together of course with descriptions of the ways in which parts are put together – roughly, biology is to physics as radio-engineering is to electromagnetism) except the occurrence of sensations seems to me to be frankly unbelievable'.[1] Sensations here of course mean sensations$_2$ and not just responding to stimuli. It is not clear why this is unbelievable – there are people who believe exactly this and think that they do so with good reason – except in the light of the assumption that everything must be explainable in terms of physics. But the proposition 'everything must be explainable in terms of physics' is not logically necessary. Yet Smart inclines to take it in this manner, or at least as an empirically well-established theory, which it is not.

The thesis that sensations are brain processes, says Smart, is not the thesis that 'after image' or 'ache' means the same as 'brain process of the sort X'. The thesis is

> . . . that insofar as 'after image' or 'ache' is a report of a process, it is a report of a process which *happens to be* a brain process. It follows that the thesis does not claim that sensation statements can be *translated* into statements about brain processes. . . . All it claims is that insofar as *a sensation statement is a report of something* that something is in fact a brain process.[2]

Here Smart makes two remarks: (1) Insofar as it is a statement about a process the process is a brain process and (2) Insofar as it is a statement about *something* that something is a brain process. I think that the first remark is acceptable while the second is not. 'Sensation' when used as a

[1] Ibid., 161.
[2] Ibid., 163.

mental concept is supposed to include the raw feel of $experience_2$. $Experience_2$ is not a process, it is a state of conscious awareness that something is the case, such as pain in the stomach, when pain can very well be the result of a brain process. There is no process involved in the conscious awareness of pain, *qua* awareness, although there is a process involved, a brain process, in the occurrence of the sensation of pain of which $sensation_2$ is an awareness. So insofar as a statement about a sensation is a statement about a process, that process is a brain process, but a statement about sensation may be about the experiencing of pain, that is, the experience of having it with emphasis on having rather than pain as a process of a distinct sort. Then the statement is about something which is not a brain process.

So the statements about sensations which are not translatable into statements about brain processes are statements about $sensation_2$. $Sensation_1$ statements seem to me perfectly translatable into brain process statements where a $sensation_1$ statement is a statement about a discriminative performance such as is involved in responding correctly to a signal which is red, or uttering the word 'pain' when in a certain state of physical disturbance. The seeing of red and the feeling of pain – these are $mental_1$ terms – such as are involved in these responses, happen because of certain brain processes and if we know what these processes are there should be no difficulty in translation. The difficulty is with regard to $mental_2$ terms where seeing or feeling is also a state of conscious awareness that red or pain.

Let us examine some of Smart's examples where a logical gap between statements does not signify different referents. One of his examples is statements about nations and individuals. Their logic is different but there is no such thing as a nation over and above individuals. But insofar as a statement about a nation cannot be translated into a statement about individuals the two statements do have different referents. Admittedly there is no entity called a nation comparable to an entity called an individual but the concept 'nation' does stand for something over and above individuals each taken singly. Individuals, as they come under the concept 'nation', are thought of as forming a collectivity or a unit for a political or some such purpose. If these different referents were not involved the logic of the two statements would not have been different in the way they are. Take again the statement 'the morning star is the evening star'. If the logic of the morning star is different from the logic of the evening star despite this identity statement, as Smart admits that they are, that is because two

distinguishable referents are involved, *the star as seen in the morning* and the *star as seen in the evening*; only it happens to be contingently the case that they are existentially the same entity. So there is a sense in which it may be said that the morning star is not the evening star, that is, if we want to emphasise the time when we see the star. If the star as seen in the morning is named '*x*' and the same star as seen in the evening '*y*' then it would obviously not do to say that *y* is seen in the morning. So despite this existential identity a statement of identity about the morning and the evening star may be false if we have two logically distinguishable referents in mind. But in the case of consciousness and brain process we do not even have this existential identity.

Take again the statement 'lightning is electrical discharge due to ionization of clouds of water vapour in the atmosphere'. 'Lightning', says Smart, is not translatable into 'electric discharge etc.' It may not be so translatable for an uninformed person who knows nothing about electrical discharge. But for someone who knows what 'electrical discharge due to etc.' means and who knows that lightning is produced by this, the two concepts should be translatable. But lots of people who know a great deal about neurophysiology – some reputed neurophysiologists, for instance – find it impossible to identify consciousness (raw feel experience) with a brain process, because there obviously appears to be a categorical distinction involved, a distinction between two orders of reality, one physical and the other non-physical, not just two different things of the same order, like lightning and electrical discharge, one of which, lightning, can be a manifestation of the other, electrical discharge, in certain specified conditions. I am not arguing that sensations cannot be brain processes on the ground that someone who knows nothing about brain processes can still talk about sensations. But if a man who is in a state of conscious awareness that red ($sensation_2$) watches the neural event representing perception of red ($sensation_1$) on a cerebroscope and finds that the two states involved are of a different order, one spatial, the other non-spatial etc., then it becomes difficult to understand what exactly is meant when one says that they are identical. Such difficulties would not arise in the case of cloud and lightning for people who know enough physics to understand how they are produced, and for them a cloud statement would be translatable into a water droplet statement, or a lightning statement into electrical discharge statement without much trouble.

Smart recognises, unlike Place, that whether consciousness is identical with a brain process or not is not an empirical issue and we

cannot therefore pin our hope of establishing this on future increase of knowledge. For no matter what neurophysiology has to tell us it will always be possible to insist that consciousness is not identical with a neural process. This will be possible, so it seems to me, because of the nature of consciousness, which unlike an electrical discharge or a performance of the organism, is present to itself – consciousness that *p* is also consciousness of itself – and is immediately felt as such. This makes the identification of consciousness with electrical discharge quite unlike the identification of morning star with evening star or of cloud with water droplets. It can hardly be denied that there is consciousness and that consciousness does appear to be something quite different from electrical discharge; if it did appear to be the same as electrical discharge – as lightning does after explanation even if it did not do so initially – there would have been no need for so much argument by identity theorists to prove their identity. Perhaps an identity theorist would say that the appearance of consciousness as something quite different from electrical discharge is a grand illusion from which we suffer and which can be got rid of by means of right knowledge – not empirical knowledge (which Smart is aware cannot settle the issue) but shall we say metaphysical knowledge – and that there is ultimately nothing but physico-chemical processes in the universe whatever else may also *appear* to be there. This will be the Indian doctrine of illusion in reverse which says that there is only consciousness in the world and the appearance of this consciousness as the physical world is a grand illusion which can be got rid of by means of right knowledge (this also is metaphysical in nature).

Richard Rorty

Richard Rorty in his article 'Mind-Body Identity, Privacy and Categories'[1] says that there are two forms of the identity thesis, the translation form and the disappearance form. To say that a sensation is identical with a brain process involves this difficulty: a sensation of pain may be stabbing but the brain process cannot be described in this way. So philosophers have criticised the identity thesis by saying that it involves a category mistake. Rorty contends that the theory can be stated in such a way that it does not involve category mistake. One may try to do this with the translation form. A statement like 'I am having a stabbing pain' could be translated into topic-neutral language of the form, 'there is something going on which is like what is going on when

[1] In *Philosophy of Mind*, ed. Stuart Hampshire (Harper & Row, New York, 1966).

etc.', and then it could be said that there would be no difficulty in identifying it with a brain process. But it is not easy to find topic-neutral language for sensations, so Rorty suggests that we adopt the disappearance form in explaining what identity means in this context. According to Rorty it is the sort of relation that holds e.g. between quantity of caloric fluid and mean kinetic energy of molecules.

> There is an obvious sense of 'same' in which what used to be called 'a quantity of caloric fluid' is *the same thing* as what is now called a *certain* mean kinetic energy of molecules, but there is no reason to think that all features truly predicted of the one may be sensibly predicted of the other.[1]

As talk about caloric fluid has disappeared with the progress of science leaving only talk about kinetic energy of molecules, so with the future progress of psychophysiology talk about sensations of pain could disappear leaving only talk about brain process. Rorty does not think that this is going to happen because of the inconvenience involved but plainly he rejects the thesis of the logical intranslatability of the language of sensation into language of brain process advanced by Feigl, Place, and Smart.

Sensation statements no doubt have a reporting function which brain statements do not have. But according to Rorty what one is really reporting when one is reporting pain is, 'my c-fibres are firing', and this will become clear when we have acquired enough knowledge of neurophysiology. Both the explanatory and reporting functions of statements about sensations can be taken over by statements about brain processes. To see this is to eliminate useless questions about sensations from neurophysiological enquiry, leaving however our ability to describe and predict undiminished. This seems to me perfectly valid, for it is much more profitable for neurophysiology to ask questions about brain processes in order to see their interconnections as well as their role in our actual responses than to ask questions about mental items. But it is at the same time true that neurophysiological questions are not the only questions to ask about human behaviour and there may be a different framework of enquiry in which questions about sensations$_2$ or consciousness may become meaningful and this may have as much relevance for a philosophy of mind as neurophysiology. Anyway according to Rorty we could convey all we convey by

[1] Ibid., 33.

'I am in pain' by saying instead 'my c-fibres are firing'. Not only that; 'my c-fibres are firing' is the information that I am really conveying when I say 'I am in pain' although I may not know this to be the case.

Some philosophers have objected to the suggestion that to say 'I am in pain', is the same thing as to say 'my c-fibres are firing', and even some identity theorists have found a logical gap between them. But it seems to me that when it does become common knowledge that pain is the result of firing of c-fibres one can very plausibly convey the fact that one is in pain by saying that one's c-fibres are firing. Also, when we do come to have this information at the commonsense level there should be no difficulty in seeing that at least one of the things I am reporting when I am saying 'I am in pain' is that my c-fibres are firing. I do not have to have a periscope hitched up to a microscope to give me a view of my skull in order that I may report a brain process, once I know through a thorough knowledge of neurophysiology which brain process stands for which sensation. Nevertheless it seems odd to say that a man who has never heard of brain processes and who is saying 'I am in pain' is really reporting here and now about his brain processes, even though it is perfectly correct to say that he would not be reporting what he is reporting unless he had the appropriate brain process, as well as that he could be made to see, by supplying him the necessary information, that there is a sense in which what he is reporting is the firing of his c-fibres. Thus it is perfectly possible to report in a sensation statement public events such as brain processes, provided both the reporter and the reported understand this language. But it does not seem to me true that one cannot possibly report anything but public entities when one reports sensations.

What then may we say a man is reporting when he is saying 'I am in pain'? He may be reporting that it is pain and not something else that he is having, such as an itch or a ticklish sensation. But he may also be reporting his *having* of the pain, meaning his conscious awareness of the experience of pain that he is having which is distinguishable from responding to pain sensations by exclamations, grimaces etc. If the man's intention is to talk about pain when he says 'I am in pain', his emphasis is on the public reference of the term 'pain' whether or not it is thought to be brain process and insofar as this is the case it is not his intention to report something private to himself. Nevertheless he may be reporting something which is in some sense private – his conscious undergoing of an experience of pain. I do not mean by using the term

'private' that we do not understand what he is talking about when he reports something which is private, it is only private in the sense that his undergoing of this experience cannot be jointly shared by another man although the latter can equally undergo his own experience of pain. It is only private in contrast to a process in his brain which can be intersubjectively looked at. If the man is doing this, if his object is to talk about his experiencing, about his being aware of pain, then he is not talking about a public object as he is doing when he is talking about pain, a $mental_1$ event which consists of firing of c-fibres and a performatory response to it. So a $mental_2$ event, the having of an experience of pain, does not disappear simply because one knows that the occurrence of pain (a $mental_1$ event) means the firing of c-fibres.

I shall in the light of the above consider Rorty's criticism of Baier's[1] 'privacy' objections to the identity thesis. According to Smart we accept Smith's introspective report about his experience in preference to what others say because at the present stage of lack of development of brain process theory we have not yet got a public criterion. When that criterion is developed we shall prefer public reports about experience to private introspection. Baier thinks that this is simply a confusion. Introspective reports are necessarily about something private and this is incompatible with having a public criterion. Baier agrees that physiological theory may so progress that we shall come to have absolutely reliable encephalograph-based claims about people's experiences. We may then know for certain that Jones is having a pain even though Jones says 'no' (he may be lying or concealing the fact of his pain for some reason). But this will never show to the sufferer that he is mistaken in thinking that he is having a pain, for such a mistake is inconceivable. Physiology can thus give us evidence but not criterion for saying that Jones is having an experience of a certain sort. For this, says Baier, the sufferer's epistemological authority is final. Hence introspection talk and brain process talk are not the same.

Rorty points out that if a physiological law connecting a brain process with sensation of pain does become well established (that is, encephalograph reports do always agree with people's reports of sensation of pain) then we are entitled to think that Jones is mistaken if the encephalograph says that he is having pain and he says 'no'. Jones, we would think, must be misjudging or misnaming; he does not know that what he is feeling is called 'pain'. According to Baier 'however good the evidence may be such a physiological theory can never be

[1] Kurt Baier, 'Smart on Sensations', in *Australasian Journal of Philosophy*, 40 (1962).

used to show to the sufferer that he was mistaken in thinking that he had a pain, for such a mistake is inconceivable'. Rorty replies

> We now see that the claim that 'such a mistake is inconceivable' is an ellipsis for the claim that a mistake, made by *one who knows what pain is*, is inconceivable, for only this expanded form will entail that when Jones and the encephalograph disagree, Jones is always right. But when formulated in this way our infallibility about our pains can be seen to be empty.[1]

Now I do not know exactly in what respect Baier thinks that mistake about experience is inconceivable. Is it inconceivable in respect of the particular experience reported (in this case pain) or in respect of the having of the experience? Baier *does not say* that if the encephalograph says 'yes' and Jones says 'no', Jones' 'no' must be taken as a proof that the encephalograph is mistaken. On the contrary he suggests that Jones may be lying. What he says is that a physiological theory can never be used to show to the sufferer *that he is mistaken in thinking that he is having a pain*, for such a mistake is inconceivable. It may be that the mistake that is inconceivable for the sufferer to make is not so much about pain but about *thinking that he is having a pain*. That is, it is inconceivable that one would think that one is having an experience x, with emphasis on having rather than x, when one is not having any experience at all. That is to say, if somebody constructs an instrument designed to detect whether one is having an experience, that is, is in a state of conscious awareness (that x) and this instrument said 'no' when Jones said that he was in a state of conscious awareness (that x) we would have to believe Jones rather than the instrument. But it is not inconceivable that one may make a mistake as regards what the experience is about. It may be in fact an experience y when one says that it is x. It is highly unlikely that such a mistake will be made about pain, but if the experience is a particular kind of pain then one having the pain may not be able to correctly identify as to what kind of pain it is. So we may agree with Wittgenstein's point that sensation reports must conform to public criteria or else be disallowed and yet claim that there is something private about the having of a sensation and that it is impossible for one to believe that one is having a sensation when one is not, although one may be mistaken as to what the sensation is sensation of.

[1] Ibid., 53.

The criteria of privacy that Baier gives seem to me to fit better the *having* of a sensation of pain, that is, conscious awareness that p rather than the sensation itself, that is, pain understood as a certain kind of physical process. For pain understood as a physical process is not imperceptible by the senses if we include the cutaneous sense, but imperceptibility by the senses is one of Baier's criteria for privacy. The others are: 1) necessarily owned, 2) necessarily exclusive and unshareable, 3) necessarily asymmetrical – it makes no sense to say I could see or hear that I had a pain but it makes good sense to say that I could see or hear that X had a pain, 4) necessarily of such a nature that the person having it needs no evidence for it, while others may, 5) it is something about which the person whose private state it is has final epistemological authority. It makes no sense to say 'I have a pain unless I am mistaken'.

If we think of the experience as not just pain but a particular sort of pain – stabbing pain, shooting pain, a throbbing pain, a dull ache or just local sensitivity – it is easy to see that the agent having the pain is not infallible or even necessarily the best judge as to what sort of pain it is, and a sufficiently refined encephalograph report may give us more correct information than the agent's introspective report can. I myself sometimes find it difficult to give an accurate description of a pain when asked by a doctor to do so, but the doctor with his long experience and with the help of an examination of the source of pain may know better how to identify the nature of the pain. Anyway, this shows that the sufferer's epistemological authority is necessarily final not with regard to the accurate nature of the particular experience had but with regard to the having of the experience. Rorty can no doubt say to this that if we are studying human behaviour then the mere information that one is having an experience, even if this is infallible, is hardly of any value. We want to know what sort of experience one is having and for this publicly inspectable states or properties, whenever obtainable, are of greater value than introspective report. This seems to me quite true. But these public properties will only explain the nature of an experience, what it is about, rather than the fact of the occurrence of a conscious having of an experience for which the agent's epistemological authority does seem final. Neurophysiology may have no use for this latter, for its interest lies in finding interconnections between different states and processes, neurological and physiological. But what is useless for neurophysiology may be useful for certain other purposes relevant to what is called 'philosophy of mind'.

THE IDENTITY OF MINDS AND MACHINES

The force of the contention that consciousness is nothing other than electrical activity in the brain which is explainable on physico-chemical laws has been strengthened by recent developments in cybernetics. As machines give evidence of functioning, in performatory terms, which if given by human beings would be taken as mental, some philosophers believe that there is no qualitative gap between machines and human beings and the same sort of laws are sufficient to explain the nature of both.

It is now widely known, as D. M. Mackay[1] says, that fundamental features of personality and mental life are linked with biochemical, electrical features of the bodily structure and activity. And as Dr Weiner stresses, there are striking resemblances between the activities of the electrical brain and the human brain as the latter's activity is reflected in recorded electrical impulses and wavelengths. According to Professor N. S. Sutherland

> We shall come to treat information-processing machines in the same way that we treat people. And we shall treat people in the way that we treat machines. The first process will be forced upon us, as the ability of computers to process information comes to match and surpass that of man. The second process will result from our increased understanding of the way in which our own behaviour is determined by the atoms and molecules by which we are composed.[2]

Professor Sutherland says that along with our progress in building artificial intelligence we are increasing our understanding of how the brain functions. 'The human brain is the most efficient information processing device known to man.' On the one hand machines help us understand how our brain functions, on the other, an understanding of how the brain functions will help us to build more efficient machines which will do virtually everything that a human being does. This sort of mutual use of information is possible because the brain itself may be said to be a highly complex computer built of course of biological material rather than of electronic valves and wheels.

It is now known that mental activity involves electrical discharge in the brain which is suffused with a steady low voltage background wave (which changes patterns with mental states); also that chemicals play a

[1] 'From Mechanism to Mind' in *Brain and Mind*, ed. J. R. Smythies (Routledge, London, 1965).
[2] *The Observer*, 9 April 1967.

part in transmitting nerve impulse as well as imprinting memory. Anyway, the principles known to be involved in the operation of computers apply equally well to the functioning of the brain itself which is a mechanism composed of complicated neural circuits into which information is fed by the senses. Such information is then processed and either acted upon or stored for further use as required. It operates just like a computer through feedback and information processing control systems and just as a machine performs a set of operations according to a set of rules so does the brain. So responses of the brain are determined by the way it is programmed, the rules according to which it is built and the information which is fed into it; there is no room for anything else such as psychical agencies.

Classical psychology supposed that human behaviour needed explanation in terms of such things as intelligence, desire, motive, purpose, choice, etc. which are mental in nature. Although mental, they bring about certain activities in the brain (they in their turn may be produced by bodily processes) which are of course physical. It was thought that as intelligence, purpose etc. are irreducibly mental the laws according to which human beings operate are psychophysical rather than merely physical. An effort was made to study these internal mental agencies introspectively in order to see how these are related to human behaviour. But not much progress was made in finding these so-called psychophysical laws and introspective study of entities believed to be mental was not very productive.

The picture has now changed and the study of human behaviour is now progressing fruitfully along a different route. Computers are now in existence which perform such highly intelligent functions as composing music, besides solving mathematical problems and playing chess, all through the functioning of physical laws. And there is no reason why yet more spectacular advances should not be made in improving the abilities of these machines. Indeed Professor Sutherland says that a time is in sight when all the most important intellectual work will be done, and the more far reaching decisions will be taken, by our artifacts with only trivial assistance from ourselves. These developments in computer mechanism have greatly stimulated behaviourist and neurophysiological studies of human beings which proceed on the assumption that a human being is a stimulus-response mechanism that works according to laws, with the result that we now have a much greater understanding of how human beings operate than in the days of introspective psychology.

There are of course people who object to the use of such terms as intelligence and purpose in respect of machines. They say that machines do what we tell them and their thoughts are mechanical, being programmed by us, whereas a human being's thought is his own. This is a confusion. If the solving of a mathematical problem involving the use of successive logical steps is to be called 'thought' in the case of human beings there is no reason why the same ability as displayed by a machine is to be denied the name whether or not the ability has been built into it. For the human ability too is there by virtue of its constitution and it can be said that if the computer has been programmed so has been the brain. It has also been said that computers do not use concepts in their thinking as human beings do, so their operations, however spectacular, cannot be said to be thought. But computers have been designed which are capable of pattern recognition, and what is use of concepts but pattern recognition. For the ability to use concepts is the ability to see different things as exemplifying a certain common form. Again it has been said that computers do not have a purpose, for their purpose is not their own. This too is a prejudice. For the activities of the self-adjusting, self-correcting type of machine is not only goal oriented, it is goal guided, and more and more complex machines can be constructed which approach closer to the type of purpose which human beings have. 'Human purposes are normally *hierarchically ordered*. . . . There is no difficulty in arranging a similar hierarchical ordering of goals in an artifact. We require only to allow higher-order feedbacks to control selective mechanisms operating on the range of lower-order goals within the scope of the artifact'.[1] At any rate whether something has a purpose or not is shown by whether it can sustain activities in a certain direction until such time that the goal is reached, correcting its moves in the process if need be. There is no doubt that machines can do this, so they can hardly be described as purposeless.

Now it has been known even before the advent of computers that a human personality would not be what it is without its possession of the brain and the nervous system and that the brain is some sort of mechanism the functioning of which does follow certain laws. Of course old fashioned psychologists believed that the laws according to which a human personality functions are two-fold, physical and psychical. That is, they believed that mental causes somehow bring out physical events which could be unrelated to antecedent physical events and which could just appear out of nowhere, as it were. This

[1] D. M. Mackay, op. cit., 165.

account does not seem plausible any longer and people who still believe that a human personality is something more than a physico-chemical mechanism because of their capability for experience$_2$, which involves consciousness, will have to say that the functioning of the brain, generation of nervous impulses, their passage and interconnections etc. are regulated according to physical laws. But as the brain is highly complex in structure and function, and is open to a wide variety of influences, it may hold the possibility of nervous activity in more than one direction at any one moment – all these activities being related to antecedent states of the neural mechanism – and possibly all that consciousness does, when it acts on the neural mechanism, is something akin to switching on current in one particular direction rather than generating it.

A mental 'cause' is then not a cause in the full sense of the term since it is not sufficient by itself to bring about the physical effect the possibility for which must already be given in the existing state of the neurophysiological organism. But of course the computer account cannot accept even this modest function of consciousness. It says that purposeful intelligent activities which apparently appear to be such that we are inclined to think that they have been brought into being by mental causes – 'mental' meaning something extra physico-chemical – can all be explained by physics and chemistry. However, I think that this is compatible with the function of consciousness in influencing the occurrence of certain events in the neurophysiological organism, for these same events could, in other circumstances, take place without the presence of consciousness on the scene – it is a contingent fact that on certain occasions certain events happen in the brain as a result of influence of consciousness on it – and being related to antecedent physical events they remain explainable on physico-chemical laws. There may be occasions when the same events take place without the agency of consciousness as an influencing factor, when the given state of the organism is such that the events in question occur as a result of being determined by certain antecedent events according to laws of physics and chemistry. The only difference is that in the case of mental influence one has to say that one of the reasons why they are there is mental in nature as far as this instance of their occurrence is concerned.

Take, for instance, a case where X decides to use his time playing chess instead of reading a novel. Whether he reads or plays his activity is made possible by the possession on his part of sense organs and a

brain where certain connections between different neural states and processes have been made by training and habit according to laws of physical functioning, and where electrical discharge of a certain sort must take place in order that he may decide between reading and playing (that is, in order that it may be a situation of choice between reading and playing and not something else). So his chess playing can be explained by neurophysiology, as could be his reading had he decided to read, although on this occasion it would be true to say that he is playing chess rather than reading because he has decided to do so. If neurophysiology could show on empirical grounds and not just on the ground of the *a priori* doctrine of determinism, that consciousness cannot activate a certain passage of nervous impulse in the brain, then of course one would have to accept that there is no qualitative difference between machines and human beings. But no matter what the advance in neurophysiology it will always be possible to insist that consciousness does influence the functioning of the brain, but this being outside the scope of the assumption and technique of neurophysiology it is not surprising that details studied by neurophysiology show no such influence. Of course if the claim that consciousness influences the functioning of the brain is not empirically falsifiable it is not empirically verifiable either, for it is always possible to explain whatever happens in the neural organism in terms of antecedent states whether or not consciousness intervenes in some fashion. So we shall have to accept whichever account, purely physical or psycho-physical, appears to us the most satisfactory, considering all the facts that we know about ourselves (including the fact of experience$_2$) in relation to the functioning of machines and the physical universe.

It has been said that either we give up the use of non-physical concepts in respect of human beings or admit that the same concepts which apply to human beings apply to machines, as the fundamental principles on which human beings perform their intelligent functions are the same as those involved in the performance of machines. This according to me is perfectly valid as long as we use the term 'mental' as mental$_1$ and not as mental$_2$, which involves conscious awareness that such and such is the case, as distinct from performance, including verbal performance, which machines are quite capable of. If a day comes when we shall have no doubt that machines possess ability for this mental$_2$ experience as we have no doubt that all human beings do, in spite of the fact that it is in some sense private, then we shall have to extend this as well to the machines.

Although there is no reason why we should not extend mental$_1$ concepts to machines most people feel some resistance to conceding that machines are capable of thinking in the human sense, no matter how brilliant their performance, and there are some reasons for this resistance. Thinking, insofar as it means the following of logical steps in a train of reasoning that lead to a definite conclusion in aid of solving a problem, is doubtless a function of which machines are capable – maybe, in some ways, they are better than human beings. But what is called thinking in the case of human beings is not confined to logical operations. Consider the question put by A to B, 'Would you be willing to leave your present job and accept one abroad with us?', to which B replies, 'I shall have to think about it'. It is this latter use of thinking which appears inapplicable to machines, which are built on a set of logical principles. This appears inapplicable not because thinking here involves logical steps too difficult for the machine to undertake but because it involves putting together a great many considerations including valuational preferences as well as the requirement that if these different forces balance themselves out the conflict has to be resolved by a decision. It is this concept of decision, to be used in a situation where there is no clear-cut solution either way and where consciously making up one's mind, possibly on insufficient evidence, is involved—that is to say, where thinking according to logical operations is not enough, that appears to be inapplicable to the activities of machines.

Of course comparable overt responses may be obtained from a machine by putting in a randomising device which will enable it to act in one fashion rather than another where opposing logical considerations balance themselves out. But the fact is that although taking a decision of this nature is not a purely logical operation, it is definitely not thought to be a random act by the agent concerned. On the contrary he feels himself the author of his decision despite the fact that the decision did not logically follow certain given data. The decision arrived at may be chance in the sense that antecedent conditions make both 'yes' and 'no' possible, but it is not chance in the sense of being fortuitous. Thus a non-logical function like decision may give direction to the actual course of logical operations in the thinking of a human personality the carrying out of which logical operation is explainable on neurophysiological laws. Machines are not persons, they do not consciously experience situations in which they act, nor are they faced with the need for this sort of decision (I am not saying

that the concept of decision defined somewhat differently is not applicable to machines).

However, the 'yes' to the question whether one would take up a job abroad may be the result of a decision in the aforementioned sense or it may be straightforward logical deduction. To the prospective employer who is interested in the man it makes no difference which it is, that is, so far as the overt verbal response is concerned it is the same (involving the same physiological mechanisms) no matter how it is arrived at. As a machine is capable of saying 'yes' as a result of logical deduction, its response overtly considered is in no way different from that of a human being whose response may be the result of decision. So I find the claim that machines are in principle capable of doing all that human beings do acceptable, although of course with the qualifications already made.

It has been said that there is at least one thing that machines cannot do which human beings can. According to Gödel's theorem, in every consistent system there is a proposition which is true but which the system itself cannot prove although we can see it to be true. We are therefore said to have abilities beyond the range of a machine. This argument no longer appears conclusive to me. The machine's ability to prove a truth about itself should not be compared with our ability – we who stand outside the system – to prove it. If we can see this truth, so can another machine which stands outside the system about which the truth is; only if we ourselves are capable of formally proving all truths about ourselves or even of knowing them all – the machine has to prove a truth about itself – can our superiority to the machine be established. I think however that we are in the same position as the machine in this respect, in as much as there are truths about ourselves which others see but we ourselves do not. There are unproved propositions in all systems including that of a human being. I do not mean by this of course that a human being is nothing but a machine, but the difference lies not in any capability of a logical or intellectual nature, it lies in the extra dimension that human personalities have by virtue of their capability for conscious awareness.

Hilary Putnam in his article 'Minds and Machines',[1] contends that the mind-body problem is not a real problem if by this is meant the problem of interaction between the physical and the non-physical (this simply does not happen). Insofar as there is a problem this can be reproduced in a Turing machine which is capable of answering

[1] *Minds and Machines*, ed. A. R. Anderson (Prentice-Hall, Englewood Cliffs, N.J., 1964).

questions about its own structure. This shows, according to Putnam, that the problem has nothing to do with human subjective experience.

> . . . We can perfectly well imagine a Turing machine which generates theories, tests them (assuming that it is possible to 'mechanize' inductive logic to some degree) and 'accepts' theories which satisfy certain criteria (e.g. predictive success). In particular, if the machine has electronic 'sense organs' which enable it to 'scan' itself while it is in operation, it may formulate theories concerning its own structure and subject them to test. Suppose the machine is in a given state (say 'state A') when, and only when Flip-Flop 36 is on. Then this statement: 'I am in state A when, and only when Flip-Flop 36 is on' may be one of the theoretical principles concerning its own structure accepted by the machine.[1]

Putnam adds, 'Now all of the usual considerations for and against mind-body identification can be paralleled by considerations for and against saying that state A is in fact *identical* with Flip-Flop 36 being on.'[2] State A is directly observable, Flip-Flop is something that the machine knows only via inference; how can they be the same? Yet it is clearly unsatisfactory to say that there are two different things. In the case of 'I am in pain' (a mental occurrence) uttered by X when in pain (a physical occurrence), mental and physical stand in similar relationships.

A Turing machine has certain logical states (rules determine the order in which these states succeed each other) as well as certain physical or structural states (non-logical) arising out of its construction. Clearly the logical states cannot possibly be there unless the physical states happen to be what they are and given the physical states the logical states would also be there. Now Putnam suggests that just as there are two different descriptions of a machine, the engineer's blueprint and the logician's machine table, so there are two different descriptions of human functioning, the physical and the mental. The mental corresponds to the logical states of the machine and physical to the structural ones. This seems to me very illuminating if mental is taken as mental$_1$, that is, capacity for response that is intelligent, discriminative, purposeful, goal-guided etc. This type of capacity can be said to be the logical counterpart of physical states which are neurological in nature. But this is not how Putnam illustrates the distinction. Instead of taking

[1] Ibid., 73–4.
[2] Ibid., 74.

the physical as neurological states and the mental as dispositional characteristics of a certain kind that issue in characteristic kinds of overt performances, corresponding to the structural states of a machine – the state of its valves etc. – and its logical state – printing 'I am in state A' when in state A according to the instruction 'print I am in state A when in state A' – Putnam takes the mental as subjective impression, that is, $mental_2$ involving conscious experience. But if mental is interpreted as conscious awareness that such and such is the case it does not seem to me that it can be explained as the logical counterpart of physical or neurological states with which they are identical, as mental, that is, dispositional characteristics which issue in a certain kind of activity, can be. The behaviourist, says Putnam, studies the human being in terms which link up with physics and chemistry as an engineer studies the machine in terms of physical states, the classical psychologist studies him in terms of mental impressions which are logical states, as it were, arising out of the same physical states.

This seems to me somewhat out of focus. The behaviourist psychologist surely does not study neurology (the structural states) or at least this is not his primary object of interest (as it is of the neurologist), he studies overt human behaviour, performances, responses etc. which of course he believes are logical outcomes of neurological states. If we want to make the mental-physical distinction correspond to the logical-structural distinction of a Turing machine the mental should be intelligent performance or disposition which is what a behaviourist psychologist actually takes the mental to be, rather than the mental impressions of the nature of private experience of classical psychologists. If the mental-physical distinction corresponds to the subject matter respectively of behaviourist psychology and neurology, rather than to behaviourist psychology and classical psychology, then of course identity does or does not obtain between them as it does or does not between the logical and structural states of a Turing machine. Putnam remarks that anyone who wishes to argue for the existence of souls on the basis of the distinction between mental and physical will have to be prepared to hug the souls of Turing machines to his philosophic bosom since this distinction in the shape of logical-structural obtains in the case of machines. But this does not apply if mental is $mental_2$, conscious experience that *p*, and not intelligent performance or response suitable to the perception of *p*. So long as we take the mental as $mental_1$ – performance, response etc. – we cannot ascribe souls to human beings without granting them equally to machines. But as machines do not have phenomenal experience consisting in a conscious awareness that *p*, ascription of souls (that is, a non-physical capacity) to human beings on this ground does not at all compel us to grant the same to machines.

Putnam says that people feel that as 'I am in pain if and only if my c-fibres are stimulated' is a synthetic proposition – being in pain and having c-fibres stimulated cannot possibly be identical. But 'I am in state A if and only if Flip-Flop 36 is on' is equally a synthetic proposition, so they should find difficulty in equating being in state A and having Flip-Flop 36 on. He concludes

> It is instructive to note that the *traditional* (italics mine) argument for dualism is not at all a conclusion from 'the raw data of direct experience' (as is shown by the fact that it applies just as well to non-sentient machines), but a highly complicated bit of reasoning which depends on (A) the reification of universals (e.g. 'properties', 'states' 'events'); and on (B) a sharp analytic-synthetic distinction.[1]

This seems to me obviously false. The traditional argument for dualism, such as that of Descartes, *was based* on the raw data of direct experience and it knew nothing of the logical-physical distinctions of Turing machines, nor did it rely on the analytic-synthetic distinction. These arguments are recent but dualism is an old doctrine which, contrary to Putnam, does base itself on the raw data of experience which is what makes the assertion of the identity of consciousness with a state of electrical discharge in the brain unacceptable and not the fact that the proposition which asserts this identity is a synthetic proposition. As I have argued earlier, a proposition like 'a brother of Y is a friend of Z' is a synthetic proposition and yet there is no difficulty in understanding that we are here talking about an identical person who is both a brother of Y and a friend of Z although being a brother of Y is logically independent of being a friend of Z.

Of course, the dualists may feel that the fact that the identity proposition about the mental and the physical has to be synthetic does show that this identity is not logically guaranteed and it must therefore be proved before it can claim validity – it does not prove itself like an analytic proposition. The dualists, believing as they do in $mental_2$ on the basis of raw data of experience, feel that the identity theorists have failed to vindicate this identity. Anyway I think that those who believe

[1] Ibid., 74–5.

in the mind-body distinction in the Cartesian fashion, where $mental_2$ is something other than a logical counterpart of a physical state, must base their opinion on the existential reality of $mental_2$ and not, at least not merely, on logical points. An argument which claims that a synthetic proposition which has two distinct concepts such as 'the mental' and 'the physical' to deal with at the start, cannot show that they are identical, is only put forward because it is thought that the mental ($mental_2$, at any rate) is distinct from the physical.

To conclude. Professor Sutherland's contention that we should treat human beings as machines and machines as human beings should be qualified if human beings are capable of $mental_2$ experience and machines not. Of course, there is nothing that human beings *do* that does not involve the mechanical operations of his neurophysiological organism and such operations are mostly carried out without any intervention from consciousness. Furthermore emergence of consciousness itself as consciousness that *p* is parasitic on certain neurophysical happenings even though it may not be produced by such happenings. But insofar as human beings are capable of $mental_2$ experience and of decision they perhaps can expect to be treated a little differently from machines in certain respects.

FREEDOM

The problem of free will arises in the context of certain features of human actions. These are like the following. (1) Most human actions are undertaken because they lead to the fulfilment of the agent's wishes, motives and so on. Some human actions however are such that the agent undertakes them not because he wants to but because he is being compelled to in some fashion. (2) Human desires and motives do not necessarily form a clear-cut consistent system so that they can all be fulfilled without clash. On the contrary, some of these wishes may oppose others in some ways and the fulfilment of one wish may require the non-fulfilment of another, at least for the time being. This necessitates choice. (3) Human situations of action are complex so that there are often alternative things one can do in the same situation as well as alternative ways of doing them. This also calls for choice. (4) All human societies hold that actions of a certain nature ought to be done

while actions of a contrary nature ought not to be done. This requires that the agent should not indulge in wishes and motives which lead to actions that ought not to be done according to certain accepted social and moral standards. It is generally thought that if the agent harbours such wishes and motives there are two different ways available for preventing them from being translated into actions: 1. the agent's own effort, 2. social pressure in the form of persuasion, blame or punishment.

Now if we take the first feature into account the question that may arise is 'Is what the man is doing the thing he wants to do?' If the answer is 'yes' he is acting on free will, if 'no', if he is doing it because he is being compelled to, then he is not acting on free will. Many determinists use the term 'free will' exclusively in this sense and here to have free will means to be able to do what one wants to do. As for free choice, 'X is free to choose *y*' means that X is not suffering from any feeling that he is being compelled to choose *y*, on the contrary he is choosing it because he wants to. With regard to ordinary actions the fact that a man is doing something generally shows that he wants to do it and there is no sense in saying that he should make an effort not to do it. With reference to actions that are socially or morally disapproved of, blame and punishment can make a man want to do things he ought to do and although this involves compulsion to begin with if the man in future does what he is required to do, even if only to avoid blame or punishment, he would still be acting freely.

There is a sense of the term 'willing' in which it may be said that a man is not willing an act at all if he is being compelled to do it contrary to his own wishes. But insofar as the man is doing it, it may also be said that he is willing it but not willing it freely. But if a man is doing something he wants to do and no pressure is being put on him of any kind at the moment of action then he is freely willing what he is doing. Thus 'X has free will' means 'X is doing what he wants to do' and 'X has no free will' means 'he is unable to do what he wants to do having to do something else instead, which he does not want to do'. This I think is a perfectly legitimate use of the term 'free will' and when used in this fashion 'free will' is certainly compatible with determinism. That is, it may well be that a man is determined in wanting to do what he is doing in the sense that given his heredity, upbringing, and circumstances it can be shown that he can only want to do what he is doing, and yet as he is not feeling that he is being constrained or compelled to want to do it either by anything external to himself or

me', when his internal state Flip Flop 72 is on, not unlike a human utterance of the same nature when his optic nerve is stimulated in a certain fashion. Kurt Baier however objects that these two utterances should be taken as evidences of same abilities in robots and humans. In the case of a robot the sensation evokes the utterance, in the case of a human being the connection is not merely causal and the utterance is not evoked by the sensation. A man utters these words because he knows that he is having this sensation. Putnam replies that a robot is just the same – the connection is not merely causal and the sensation does not just evoke the utterance. The robot has to be in an internal state which realises the predicate 'knows that *p*' in order that it may utter a 'psychological' statement 'It looks etc.' Baier however thinks that there is a difference between this robot utterance and my utterance which I utter because I know that I am having a sensation which implies that the robot does not know this.

Putnam makes short work of this objection by interpreting 'knowing that' purely as being able to undergo a process of sensation or behave in a certain way. If knowing that one is having a sensation is identical with having a sensation then the robot too knows that. If knowing that consists in being able to describe then the robot knows that too. The psychological-physical distinction is applicable to robots if it is applicable to humans. A robot can say 'It looks etc.' when it does not know that Flip Flop 72 is on or even the existence of Flip Flops. The robot can even say 'I know that I am having a sensation'. These utterances are instances of differences between mental and physical states and this distinction applies to the functioning of robots in the same way that it applies to that of humans.

I agree with Putnam that both a human being and a robot will have to be in an internal state which realises the predicate 'knows that *p*' in order to utter a psychological sentence and that the distinction between mental and physical is applicable to both. It also seems to me that a human being can utter a psychological sentence of the form 'It looks etc.' in the same way as a robot without having to know that *p* in any sense that is special to him (that is, involving more than an internal state that realises the predicate 'knows that *p*'). It is not necessarily the case that I utter this *because* I know that *p*, if my knowing that is something over and above my having a sensation, and my utterance may be an exact parallel of that of the robot. But it equally seems to me that it is possible that on some occasions I utter this because I know that I am having a sensation in a sense in which a robot, as currently conceived, does not. Here I am aware of being conscious that I am having a sensation, and I suspect this is what Baier has in mind when he objects that the robot does not know that *p*. We do not believe that robots experience this sort of thing although there is a perfectly good sense – as expounded by Putnam – in which a robot may be said to know that *p*.

Putnam may say that I am making this distinction because I am prejudiced against robots. But hardware does not worry me and for my part I am perfectly willing to concede that robots can have self-consciousness in the full human sense as soon as I come to believe that I have evidence that they have it. But I need for this purpose more than intelligent performance of the type robots currently exhibit. I manage to have an intuitive understanding of human beings as humans who have private experience and I do not have to judge this by watching their performances all the time. Nor do I deduce this understanding about human beings from their appearance and body material although I associate a certain appearance and a certain body material with being human. There is no reason why this sort of understanding should not operate just as well in the case of robots. Of course, as I said, if a human being is conceived as consciousness reflected in a highly organised material system with which it interacts then there is no reason why it should not in principle be possible to artificially create a conscious being by creating a material system of the requisite degree of complexity of matter and form. As to whether we shall ever manage to accomplish this, that is another matter. But I just do not think that robots as currently conceived have self-consciousness, so they do not yet qualify for being human.

VEDĀNTA

The monistic Vedānta system of Samkara, henceforward to be referred to just as Vedānta, says that in the ultimate analysis there is no such thing as an independent material principle for consciousness to interact with. It is we who through ignorance impose the idea of an inert material world and its manifold distinctions on to the true nature of things which is pure consciousness. This material world of course exists as long as this state of ignorance exists and at this level some sort

of interaction appears to take place. But this interaction presents no problems, for the material world is really a superstructure built on the ultimate reality of consciousness through imposition and this we can see to be the case in transcendental experience. The mention of transcendental experience as a valid source of drawing conclusions about what there is would no doubt make this philosophy esoteric for Western taste. In classical Indian culture however transcendental experience was treated as something which is attainable by effort and it was commonly accepted that people did attain it. Its experiential value was thus not suspect. Anyway this philosophy has certain things to say on the nature of consciousness even if we reject the claim that transcendental experience, supposing it to be a fact, is of any philosophical consequence.

In some Vedānta literature it is suggested that consciousness is literally the stuff out of which the visible universe is produced and what is called matter is just consciousness itself which appears to have taken the shape of a non-conscious world through ignorance. This operates through imposing distinct names and forms on the undifferentiated mass of consciousness in order that it may get particularised in the way we see it in our empirical experience, particularisation being necessary for the purposes of practical living. This doctrine no doubt sounds absurd but it seems to me that it is not much more absurd than the theory that consciousness is nothing but a modification of matter or that it does not even exist as is sometimes suggested by materialists, if of course we start with the presumption that matter and consciousness are two different orders of realities as most people do because of manifest differences between the two. These differences, according to the monist of either variety, materialist or idealist, arise out of a grand illusion which we suffer as a result of ignorance and from which we can be saved by a correct theory. Of course the materialist has this advantage that he can point to laws of physics and chemistry and show exactly how matter functions, although he cannot show precisely how these laws generate consciousness out of unconsciousness in the way he can explain the generation of steam out of water through heat and molecular motion. But the idealist monist can easily concede, and he does, that consciousness, when it appears as matter (he cannot very easily deny the appearance of matter just as the materialist cannot deny the appearance of consciousness), does behave according to strict laws which we know as laws of physics and chemistry. Vedānta, unlike idealist Buddhism, does not deny the reality of matter at the empirical plane of experience, all it denies is that it is, in the ultimate analysis, something radically different from consciousness as we normally believe, and illusion lies precisely in thinking of it as something different.

Implausible as this Vedānta doctrine may appear Vedānta has an even more implausible theory to explain the appearance of matter. Pure consciousness is the only reality ultimately speaking – to be ultimately real is to be at all times, past, present and future and nothing which is not that is ultimately real – but there is another rather mysterious ontological force in the world called 'ignorance' (literally non-knowledge). This is beginningless like consciousness but unlike consciousness it disappears on the attainment of knowledge. Thus ignorance has an indescribable status and it cannot be categorised either as being or non-being. It is being insofar as it exists at all and it is non-being insofar as it does not exist any more on the attainment of knowledge. It is through the agency of this ontological force that one imposes material modifications on pure consciousness, that is, applies various names and forms by means of which differentiations are made out of something which is in fact undifferentiated. As long as ignorance exists the material world exists in actual fact on the empirical plane of experience and it must not be thought to be just a figment of one's imagination.

Indeed according to Vedānta even the object of an ordinary illusion can truly be said to exist within its own term of reference which is that of seeming. But of course the sense in which it exists is not the same as the sense in which the object of veridical perception exists. This last again although real within its term of reference, namely, utilisability which an object of illusion lacks, is not real in the sense in which what is revealed in transcendental experience is real, since pure consciousness, the object of transcendental experience, exists timelessly and an empirical object does not do so. 'Existence' is defined in this system not as a potential object of sense perception, but as being an item of experience which is acceptable on the basis of such experience as long as it is not contradicted by another experience. An illusory object exists as long as it is not contradicted by veridical perception. An object of veridical perception in its turn exists as long as it is not contradicted by transcendental experience. But the object of transcendental experience, being the very function of self-consciousness, cannot ever be contradicted. So it exists timelessly or at all times. Vedānta thus postulates a doctrine of different levels of truth and reality

and its denial of a certain level of experience must be understood to be from another and higher level. A man who is actually living within the sphere of ignorance has thus no reason to treat the world as anything but fully real at his level of experience and it remains real so long as it is not contradicted. It is thus a mistake to label the world as illusory in any ordinary sense. The material world exists as a utilisable complex and its reality lies in its being useful for the purposes of practical living. It is then true or real at its own level which is defined by practice and use. But we must go beyond this level to what is called the transcendental level and there, since its term of reference is timeless existence of pure consciousness, we cannot claim any more that the empirical world is real. The relation of ignorance to pure knowledge is incomprehensible and we cannot tell why ignorance should be there or operate by creating the empirical world through imposing names and forms on pure consciousness.

Now one may wonder why anyone should go to the length of postulating an incomprehensible ontological force called 'ignorance' for an explanation of empirical facts of experience. There are two reasons for doing so. (1) The Upanisads had declared that reality is one and that this reality is pure consciousness. This of course does not necessitate that the presence of matter should be explained in the above fashion. For there are many Indian philosophers who accepted the Upanisadic pronouncements and yet did not feel compelled to deny the reality of matter, either in the proximate or in the ultimate analysis. (2) There is the fact of transcendental experience which, according to those who had experienced it, is a heightened state of consciousness of being where being is identical with the being of consciousness. In this state the differentiated empirical world no longer exists as an object of one's experience. To Western philosophers it may appear that even if transcendental experience is a fact, and in all probability it signifies nothing but a state of physical or psychic disturbance, philosophy has nothing to do with it and philosophical analysis of what we say about the world need take no notice of it. This obviously is not the attitude of the Indian philosophers we are discussing, possibly because for them transcendental experience was a culturally accepted fact and as such respectable. Anyway, ancient Indian culture being pre-scientific, explanations in terms of physical and mental disturbances were not forthcoming, and those who took it seriously felt that reality must be explained in such a way that it fits this fact. Hence the doctrine of different levels of reality.

The explanation put in commonsense terms is like this. We are in a state of ignorance about the nature of consciousness when we believe that consciousness simply means empirical consciousness of this or that and we remain in this state of ignorance as long as we do not have transcendental experience. This experience is experience of consciousness as pure consciousness, devoid of all empirical connotations of a particular nature. We see ultimate truth when we see reality like this and when we see reality like this we no longer see the empirical world. This simple explanation then gets hypostatised into a doctrine which posits 'ignorance' as an ontological force on its own and then relies on an incomprehensible relation of this force to reality for an explanation of the world.

The fact is that even if we accept that transcendental experience is a fact and not just a pathological condition there is no reason why we should accept this experience as the only experience which has any claim to ultimate validity. If we define ultimate reality as that which exists timelessly or at all times, it follows from this definition that the world is not ultimately real. But this only means that the world does not exist timelessly. Those who claim the reality of the world do not claim that it exists in this fashion, so there is really no argument here. As for empirical experience being contradicted by transcendental experience inasmuch as the world ceases to be there at this level, the contradiction is only apparent. For one experience to contradict another they must be about the same thing. But transcendental and empirical experience are not about the same thing. Transcendental experience is an experience which specifically concerns the nature of consciousness as a function of self-manifestation exclusive of its function as other-manifestation. But it is this other-manifesting function of consciousness which is prominent in empirical experience because we are here concerned not with pure consciousness but with objects. No wonder that these two experiences do not reveal the same things and since they are not about the same things there is no contradiction. There is then no reason why the level of experience called 'empirical' should not be taken as just as revelatory of what there is as transcendental experience. It is only when one holds on to a monistic doctrine at all costs that one feels tempted to deny the ultimate or genuine validity of one or the other and consequently the reality of what is revealed therein. The idealist monist rejects empirical experience as not ultimately valid, the materialist monist refuses to accept that the supposed experience of consciousness as a pure function of

self-manifestation is a genuine experience at all (it must be induced by abnormal conditions in the organism) which can claim any objective validity.

It seems to me that both the idealist doctrine that our empirical experiences are not ultimately real and the materialist doctrine that there cannot be any such thing as experience of pure consciousness are not so much factual assertions that can claim experiential validity as concealed value judgments. The idealist believes that the important thing about man is his consciousness which betokens a non-material principle working within him, all else being unimportant (unreal?). The materialist who accepts sense perception as the only acceptable proof of existence takes man as a highly complex physical organism which can be scientifically studied. Anything which suggests the presence of things not amenable to physico-chemical laws it considers an undue obstacle to the progress of science and so declares that it is non-existent.

Vedānta of course gives reasons why we should not accept matter as an independent category of existence just as the materialist does to show why consciousness should be conceived as just a function of the brain. The physical world, it says, is not self-manifest and it depends on consciousness for its manifestation. It therefore has no independent reality. This of course is a confusion. The physical world's not being self-manifest means that it has no knowledge that it exists and there must be consciousness in order that there may be this knowledge. But this only means that the physical world depends on consciousness for the knowledge that it exists, not for its existence as such. Why cannot the physical world continue to exist when it is not known to anyone that it exists? It does not have to be known in order to exist, although it has to exist in order to be known. All that would happen when it is not known is that no statement to the effect that it exists can be made. Now Vedānta also claims that only the self-existent can be said to exist in any ultimate sense. That which is self-existent cannot be understood as derivative of something else. The world of objects has to be thought of as derivative since an object can exist only as the effect of something else. Thus the world of objects, not being self-existent, cannot be said to exist in any ultimate sense. However, it is only if we take matter as solid inert substance in the way Vedānta does that matter cannot be thought of as self-existent. If we think of matter as the manifestation of a self-activating principle as Sāmkhya does then it cannot be said that it is not self-existent.

Another argument of Vedānta is that the real is something which is necessarily true, that is, it is logically impossible to deny its existence in a statement. It will be a contradiction to say that consciousness does not exist because the making of this statement involves consciousness (at least in the case of man) but physical realities can be contradicted in principle. Yet there is absolutely no reason why we should say that that which is not logically necessary does not in actual fact exist and matter may very well exist as a contingent fact. To say that the existence of matter can in principle be contradicted is not to show that it is in fact contradicted. If unreality simply means that the possibility of being contradicted is not in principle ruled out without this abstract possibility having to be actual in any particular case then matter may be unreal in this technical sense but this will not mean that it does not in actual fact exist.

Another sense in which matter is said to be unreal is this. The principle of consciousness as a function of manifestation is not, according to Vedānta, produced by anything else, nor can it be destroyed, for creation and destruction are categories which apply only to physical objects that have magnitude. It is then true of consciousness function that there never was a time when it was not, nor will there be a time when it will not be. But this is not true of the world. Vedānta arguments may apply only to the manifested world, but implicit in the manifested world there is a self-activating principle of materiality which according to Sāmkhya is neither produced nor can it be destroyed. So consciousness is not the only reality even in this special sense of being real for all times. Furthermore there is no reason why we cannot call something fully or ultimately real which exists for a limited period of time only, when in actual fact we do use the term in this way. What Vedānta is doing then is recommending a certain technical use for the term 'real'. But matters of fact cannot be decided by redefinition. Followers of Vedānta try to meet this argument by saying that matter is real enough within the world of use but not ultimately. But this sort of talk is needlessly confusing, for having said that matter is unreal in a technical sense one easily slips back into thinking that matter is unreal in the ordinary sense and examples of this abound in Vedānta literature. Besides, Sāmkhya claims that matter is real even ultimately, and Vedānta has not succeeded in showing that Sāmkhya is mistaken in this.

Having rejected the metaphysical doctrine of idealist monism which denies the ultimate validity of matter we may now look at the

Vedānta analysis of consciousness which makes certain important observations towards establishing the reality of consciousness as a fact *sui generis*.

The most important distinction Vedānta makes in the word of ordinary experience (which exists in some sense) is between the conscious and the unconscious, that is, between a capacity which exists as self-manifested whenever it exists, and something which is lacking in this capacity. Consciousness is defined as a function which is aware that it is, that is, as a function of self-manifestation. A thing or process is called unconscious when it is not aware that it is and needs the function of consciousness to be manifested as such and such. Vedānta rejects the idea that consciousness can be generated out of something which itself is intrinsically unconscious. It would say that the cause cannot be less than the effect. The way physical causality is conceived it ought to be possible to reduce the effect to the cause but consciousness cannot be reduced to unconsciousness for it is by definition aware of itself while unconsciousness is not so aware.

According to this definition of consciousness the mental – when all it denotes are capacities for intelligent functioning – belongs to the unconscious. Mental activities can go on on an unconscious plane and it is only when consciousness is reflected in them that they become illumined or manifested as such and such. Vedānta for the most part accepts the Sāmkhya doctrine as to how a human personality functions. There is contact between objects and sense organs, not the physical organs but capacities inherent in them, and these capacities are of the same order as capacities in objects to produce sensations in us. Through this contact the internal sense, in which mind is included, takes the form of the object it is discriminating. When consciousness is reflected in the internal sense there is conscious awareness of the object. Some Vedāntins put forward a different theory called limitation theory, according to which the function of the internal and external senses is to limit the manifesting capacity of consciousness to objects with which they are in contact. Normally what is called consciousness in men (as distinct from pure consciousness or consciousness as such) is then a manifesting function as modified by the activities of the senses. This is sometimes referred to as 'modal consciousness' to distinguish it from consciousness function as such, conceived apart from the limitations or modifications produced by the senses, internal and external.

Consciousness is defined as that which necessarily manifests or reveals itself as a result of which there cannot be any doubt about its

existence. Now Armstrong in his *The Materialist Theory of Mind*, attacks the indubitability thesis as dubious, and he takes this to be the doctrine that we have indubitable or incorrigible knowledge concerning our mental states. He also calls it Cartesian. Descartes however did not claim indubitability about any odd mental state. On the contrary only those mental states which involve a clear and distinct idea are valid according to Descartes. What is indubitable is not a mental state as we think of it, but the bare fact that there is consciousness. What he said was that when he tried to doubt that he was doubting he found that thinking was going on even while he deliberately engaged himself in doubting its existence. It is this fact of thinking or consciousness going on which is indubitable, because one cannot deny it without contradiction, not what one thinks about, nor the details of psychic states or processes. If anyone asked Descartes 'Tell me what actually is your thinking process like, as it is going on in your mind' and Descartes made an introspective report, he could very well be mistaken about some of the things he said were going on. But Descartes made no claim that mistakes of this nature are logically impossible. What is logically impossible is to be mistaken about the fact that one is in a state of thinking or being conscious whilst one is actually in this state, and it makes no difference what the thought consists of – whether one is thinking of a rose or a book – so far as this aspect of the matter is concerned. Descartes' argument was offered not as a proof of the indubitability of all mental states – only the clear and distinct ones are acceptable – but as proof of the existence of mind or consciousness as an independent substance which proves itself, according to him, even while one is doubting or denying it. What Descartes is proving here is what is called the 'self-manifest' nature of consciousness in Indian philosophy. Of course it would perhaps have been better if Descartes did not say 'therefore I exist' if by 'I' here is meant a personal ego, identical, permanent etc. He could have said 'therefore consciousness exists', and this is the sense in which Husserl and Sartre interpret the 'cogito', that is, it was not specifically a claim about Descartes himself or any of his mental states but about the function of consciousness which established itself as self-evidently given.

Vedānta uses the Cartesian argument without the personal 'I'. Vedānta would say that the fact that it was Descartes rather than Smith who went through the cogito process makes no difference to the demonstration, for the same thing about the indubitability of consciousness as a characteristic kind of fact would be revealed no matter who is

trying to deny consciousness. Indeed it claims that awareness of itself on the part of consciousness is a process in which the thought of a personal ego is not involved. It is like Hume saying that one can find a perception going on without finding an 'I' and also like Sartre's claim that the pre-reflective cogito is an impersonal process. Hume of course was not trying to establish the fact of consciousness even apart from any object of consciousness, and Sartre would insist that the ego, even if it is a construct, is real enough, unlike Vedānta which says that consciousness conceived as a function of self-manifestation does not need an object in order to be what it is and that the ego is unreal in some sense of 'unreal'. Hume of course had to say 'I cannot find an "I"' and it may appear to be a contradictory statement. But the contradiction disappears if the first 'I' stands for not a personal self but Self, Husserl's transcendental subjectivity or Sartre's pre-personal cogito. 'I cannot find an "I"' then means that a conscious process in search of a personal self fails to discover one as given to consciousness. Vedānta would say that the first 'I' is just consciousness manifesting itself as consciousness, not as an object but as a function of awareness that is aware of itself, no matter who is conducting the search. This is what is called by Vedānta 'pure consciousness'.

This consciousness was in Husserl and Sartre a logically distinguishable principle of functioning to be found in the manner of being and knowing of men – transcendental means for Husserl that which must be logically presupposed, not something that as a matter of fact can exist in isolation from all empirical psychological contexts. Vedānta claims that this function of self-manifestation is not only logically presupposed, it has also a factually distinct existence as a principle of functioning which can be realised if we learn to separate modal consciousness from pure consciousness. This principle does not need an empirical ego or a psychological self for self-manifestation which is its essential nature – although this may appear to be so in case of modal consciousness – nor does it need an object of knowledge in order to be the kind of function it is. This self-manifesting function of course functions in empirical man in association with a mind-body system – by which it gets limited in this context – in order that he may have the kind of experience of the world that he has, effected through the other-manifesting function of consciousness. But the nature of consciousness as an other-revealing or self-revealing function is derived neither from man's psychophysical constitution nor from the object it reveals. If so, it is a characteristic kind of ontological fact.

However, realist philosophers in India objected to the idea of pure consciousness existing by itself and according to them such a consciousness is a fiction. What we do find in knowledge (the term 'knowledge' is often used synonymously with 'consciousness') is a threefold distinction between the knower, the known and knowledge; knowledge by itself without a knower or something known does not exist. Some complicated arguments were used in Vedānta literature to counter this objection. The nature of these arguments show that Vedānta is using *a priori* or transcendental deduction, without using this name of course, against the realists who hold a purely psychological and empirical view of knowledge. How do you know, it asks, these distinctions, say, between the knower and the known which you say is necessary for knowledge? What is it that reveals the relation that you say knowledge must have to the knower and the known? Surely it must be knowledge itself, the manifesting power of which could not have been dependent on the objects revealed, the particular knower involved, or its relation to them. Is the knowledge of the pot different from the pot? If so, how do you know that they are different? Surely not by another process of knowledge, for in that case we would be involved in an infinite regress. And what precisely is the distinction between knower and knowledge? It is not a distinction that we are aware of at the moment when knowledge takes place. For instance, if someone says 'I know now what you mean', he is not reporting that a particular subject X has now come to possess a piece of knowledge *y*, he is only reporting that some understanding which was previously lacking has now taken place. A knower is needed, it seems, to support knowledge. If it is distinct from knowledge it must be known apart from knowledge. If it is known apart from knowledge then it is not known as a support of knowledge. If the knower himself is known by knowledge – and how otherwise can he be known – he cannot himself be a support of knowledge. To say this is to argue in a circle. The upshot of all this is that Vedānta claims that knowledge or consciousness is logically prior to the subject-object distinction which can only be revealed by consciousness. It cannot then need this distinction in order to be the revealing function that it is.

The essence of consciousness, its self-luminous character, lies, according to Vedānta, neither in the fact that there is an empirical ego who is conscious, nor in the fact that there is an object which it reveals, but in its very nature as a function of self-manifestation. This function is called pure consciousness which presupposes both the concepts of

subject and object and it is understanding of this which makes possible the use of such terms as 'direct' or 'immediate'. It is difficult to decide what the personal or empirical self consists of but consciousness grasps itself directly and immediately as itself. It is of course possible to know consciousness as an object of knowledge, that is, when it is conceived as a state of consciousness. But this so-called state of consciousness which is thus known is called modal or empirical consciousness which is simply an awareness of a mental state illumined by consciousness. And this last consciousness is not known in this fashion, it is what witnesses the other and such a witnessing consciousness is always involved whenever we know anything including our own mental states. People who say that consciousness cannot be known by itself and that it must belong to a personal ego have in mind only modal consciousness which is an illumined psychological state involving a modification of the mind (according to some object perceived or thought of) and this does belong to an ego or psychophysical personality. But the witnessing consciousness does not need a psychophysical process or a mental modification to know itself and it is never known either as a subject or object of knowledge.

Witnessing consciousness is always there insofar as even modal consciousness, which is consciousness of an object, is aware of itself, although insofar as consciousness is modal it is limited by the psychophysical organism in the context of which it functions. At the level of empirical experience therefore the witnessing function of consciousness is never known except in the context of an empirical ego's awareness of some object. When we are conscious of an object the contact of the sense with the object and the attention given to it by the mind are certainly necessary for the experience of the object to take place, but the consciousness that we are in a state of awareness regarding something does not need any extra process, any activity of the mind or the senses, to be there, it manifests itself through being what it is. Vedānta would have to say that it is just a contingent fact that in the sphere of empirical experience consciousness' awareness of itself is revealed in the context of its awareness of an object, but this other-revealing function of consciousness is not something that belongs to the very nature of the function in such a way that it cannot be what it is without revealing an object, as the realists contend. Indeed a transcendental experience consists precisely in the grasping of this self-revealing function of consciousness without that of object-revelation.

Consciousness reveals itself in the manner it does in a human personality because the human personality is an ego and the ego is a psychophysical system which limits the function of consciousness by making it predominantly an other-revealing function. But insofar as man can ask questions about his ego, the ego too is presented to a witnessing consciousness. Man is thus at once both a witnessing self and an empirical self and he is an empirical self insofar as the function of consciousness is modified in him by the limiting activities of his mind and senses into a modal consciousness which is what he ordinarily believes his consciousness to be. Man is obsessed by his empirical self as a result of ignorance and this veils his nature as a witnessing self. But this witnessing self is there whether he recognises it or not. The ego perceives only that to which his psychophysical activities are directed, but knowledge that *p* is not produced by these activities – these only produce an unconscious awareness of *p* – and it needs a witnessing self for this knowledge to be there.

This idea about the limiting function of the mind and the senses has some similarity with Bergson's contention that the brain is an organ of attention à la vie. What it does is to cause men to perceive a limited part of their environment which is divided up by the intellect into separate objects through the use of concepts. In Indian language the intellect through ignorance uses names and forms to differentiate the undifferentiated and a man's consciousness reveals only those objects to which his senses and mind are directed. Bergson's idea that the brain limits man's consciousness in order that he may attend to what is biologically useful is also similar to the Vedānta view that the empirical world is a world that has its value and its own level of reality as a utilisable complex useful for the purposes of practical living.

How then does this consciousness function as a self-manifesting activity come about? It does not come about, it is always there and does not have to be produced. This does not mean that we shall find consciousness to be there however far back we go to the past or further forward to the future. Consciousness function is not a temporal fact and whether it was there in the past or will be there in the future are meaningless questions. So is the question whether it is earlier or later than some other fact. So we must comprehend it as eternal or timeless. Nor is this consciousness function limited to the psychophysical system, it is all pervading says Vedānta. This should not be taken to mean that consciousness occupies all space since it is not a spatial entity at all. To say that it is all pervading is to say that it cannot be said to be limited to here or there and that it can have no particular spatial

location. This idea of consciousness being there not as a part of us as a psychophysical system will, no doubt, appear strange to many, but there is nothing particularly strange about it. Our body is peculiar to us and yet our body and nervous system are composed of elements which are not peculiar to us and the same physico-chemical functioning which our body exhibits is manifested throughout the universe. Consciousness in the same way may not be something peculiar to us as individuals so long as we think of it as a certain principle of functioning and not as modal consciousness which is limited by the psychophysical organism that is peculiar to us. So it may be said that just as our body is peculiar to us but not physico-chemical energy of which it is a particular manifestation, so our modal consciousness is peculiar to us but not the principle of consciousness of which our modal consciousness is a particular manifestation. And it is to this principle of consciousness that the categories of space and time do not apply, not to modal consciousness.

We have so far considered only logical reasons why Vedānta thinks that we must postulate a pure consciousness function. There are of course other reasons than purely logical ones why one should accept such a pure consciousness as an ontological fact. One of these is transcendental experience already mentioned wherein this pure consciousness is supposed to be actually experienced. This experience of course is not a thing of common occurrence. But it is possible, according to Yoga psychology, to achieve a state of concentration by rigorously training oneself to reject all external stimuli (by refusing to interpret them as stimuli of such and such nature) as well as to inhibit all flow of thought in the mind. This may ultimately lead to a state of meditation where one becomes dead to all stimuli, external and internal, but still remains in a state of awareness. This is a state of awareness of being but the being involved here is said to be nothing other than the being of consciousness which is no longer a consciousness of anything in particular, including the ego, since all appearances are inhibited through the stoppage of mental modifications that are responsible for empirical experience. Admittedly this is not a very easy state to achieve, for a human personality is constantly subjected to internal and external stimuli of all kinds, but it is said to be attainable with help of techniques of dissociation preached by Yoga philosophy. This is a state in which one experiences what is called pure consciousness which is said to be a state of awareness of being, but the being one is aware of is the being of consciousness which in its turn consists entirely in an awareness of well-being. So pure consciousness is defined as being, consciousness and bliss. Vedānta philosophers (and others) claim that whether or not such a state of realisation is attainable is not a matter of *a priori* judgment but experiential verification.

But there are other reasons as well. Leaving aside transcendental consciousness we have three kinds of experience of consciousness, (i) in our waking state, (ii) in the state of dream and (iii) in dreamless sleep. In the waking state consciousness is associated with the functioning of the mind via the senses. Here the externally given empirical world is predominant and consciousness which is aware of it is called modal consciousness. This modal consciousness may be explained as an occurrence that takes place through the impact of objects on the senses and the mind (brain). This is the model that realists and materialists use in their discussion of consciousness. In the state of dream there are no objects given to the senses, but here the mind fabricates images out of impressions left by previous experiences. So this may still be explained on essentially the same model. In the state of dreamless sleep even the mind is inactive and here we are tempted to say that a person in a state of dreamless sleep is totally unconscious. Vedānta however denies this. If we define consciousness as modal consciousness then of course it is true to say that one is unconscious in a state of dreamless sleep which simply means that one is not aware of anything internal or external. This is understandable, for knowledge of objects, whether of external objects or of our internal states, is possible through the modification of the mind and the senses which are inactive in a state of dreamless sleep. But it cannot be said, according to Vedānta, that all functioning of the principle of consciousness altogether stops in a state of dreamless sleep.

A man who wakes up from a state of dreamless sleep may say 'I slept soundly and was not aware of anything'. Now if we ask this man, 'how do you know that you were not aware of anything' he might be tempted to say 'because I do not remember knowing anything'. But we often do not remember our dreams, nothing at all about them, and yet know that we had them. How then do we make this distinction between not having any dreams – representing a state of undisturbed sound sleep of which we are aware as is shown by our talk about it on waking up – and having dreams that we do not remember? Let us look at the statement again. The man says 'I do not remember having any experience'. But one only remembers what one has experienced. How can one not remember an experience one never had? The statement

made on waking up from a state of dreamless sleep indicates a state of total unknowing, it is not a statement about some one thing one says one has not known because one has no memory of it. This state of total unknowing must have been felt or experienced for one to remember or talk about it.

It may so appear that the statement made about not knowing anything is a statement about the waking state, about not remembering anything at this moment. It is however intended as a statement about the sleep state and not about the moment of waking up or any subsequent state. The original statement was 'I was unaware of anything' and not 'I do not remember having known anything'. I think that not knowing anything in a state of profound and dreamless sleep has quite a different feel about it from not just remembering anything. In any ordinary use of not remembering anything – which one presumes has happened or been experienced – it is possible for one to try to remember it as one tries to remember a dream one has forgotten; but one does not try to remember what happened in a state of profound and dreamless sleep for one knows that nothing happened, that it was a real holiday from knowing. One does not infer this from not remembering, one is actually aware of not having known anything.

This would no doubt appear absurd to someone who thinks that one can only know things that are there, as for things that are not there it is really a matter of passing a negative judgment which merely denies a positive judgment that they are there. One cannot know directly that there is nothing there as one knows that there is something which is actually there. Sartre however claims that one intuits or apprehends absence directly in the same way as presence, and this one can do because of the transcendence function of consciousness. If I am expecting someone somewhere and do not find him there I may say 'I see that he is not here', and mean what I say. My grasp of his absence is direct and I am not just contradicting a statement that he is here. This is possible because consciousness as a function of transcendence can take non-being in its stride as well as being. It would no doubt be said that when I say, 'Pierre is not here' I am really aware of the fact that he is not here while in a state of dreamless sleep I am no doubt aware of nothing but I am not aware that there is nothing as I am aware that there is no Pierre here. This Vedānta would say is owing to the inactivity of mind in the state of dreamless sleep. To be aware that I am not aware of something a mental modification is necessary as actually takes place in the case of absence of Pierre. But there is no mental modification of any

kind in a state of dreamless sleep and hence no awareness of a distinct mental state constituted of not knowing anything. But the very fact that one says 'I did not know anything' and intends by this statement to refer to the state of dreamless sleep itself shows, according to Vedānta, that the function of consciousness as pure consciousness was present in this state but in a state of empirical unknowing because of the absence of the particularisation function of the mind and the senses.

This experience of course is not comparable to transcendental experience where the nature of pure consciousness is fully revealed. Anyway experience of pure consciousness is not an experience of it either as a subject or object of knowledge, it being something that is always presupposed and known only as presupposed and not as something that one can subject to one's mental gaze. It is only modal consciousness which can be witnessed, not the witnessing consciousness itself and in dreamless sleep there is only the witnessing consciousness. This witnessing consciousness remains in a state of unawareness of anything during dreamless sleep and this is what is remembered on waking up. Hence the sense of continuity despite the gap in experience and the ability of the empirical self to pick up the thread of its life across this gap without any feeling of strangeness. This also explains one's sense of propriety in referring to the state of dreamless sleep as 'I'. Without this presence of the witnessing consciousness a continuing sense of 'I' which is not disturbed by the gap created by dreamless sleep wherein there is no experience, by which one defines one's sense of 'I' in empirical living, would be inexplicable, according to Vedānta. Dreamless sleep however is still a state of ignorance hence witnessing consciousness' witnessing of itself is not experienced in the same way as in a state of enlightenment and there is only an indication of it which is also shown in the feeling of well-being that dreamless sleep produces.

Now for the problem of freedom. Vedānta doctrine of freedom is very similar to that of Sāmkhya in that both these systems are concerned to secure for man a state of unconditioned being, free from the limitations that ordinarily go with being human. But there is an important difference between Vedānta and Sāmkhya as to what this state of freedom is like and why man should seek it. Sāmkhya says that consciousness is distinct from matter and as man is consciousness he must secure his independence of matter on which he falsely believes himself to be dependent. This state of independence is conceived negatively as a state of detachment in which a man is no longer moved by happenings in his mind, body or environment and as a result he is no

longer subject to the afflictions and sufferings that being dependent on matter brings. Vedānta claims that freedom is a state of positive experience, that of being, consciousness and bliss and not merely one of absence of suffering. Being is here only the being of consciousness as such and not anything extraneous to it like the empirical object or the subject. The empirical subject has here transcended his empirical limitations by being identical with his essence and this is freedom. In this state of freedom man finds that only consciousness is ultimately real (where 'real' is really a value term which means worthwhile).

But the Vedānta concept of freedom has its difficulties. In order to emphasise the aspect of freedom of consciousness which lies in its manifesting or witnessing capacity – an activity which remains invariable throughout differences of states it witnesses – Vedānta postulates that pure consciousness is just an impersonal cogniser, not a doer or enjoyer of fruits of actions. Activity or enjoyment belong to the empirical ego characterised by the possession of mind and sense organs but the Self in man remains a mere witness capable of withdrawing itself to its unconditioned status. But this view runs into difficulties. Vedānta, unlike Sāmkhya, does not believe that the material principle is self-activating and it has to say that it is the Self which causes the mind and the body to act. As has been observed by one Vedāntic writer, 'Being devoid of the light of consciousness they cannot move of themselves'. But we cannot have it both ways – a mere observer status for consciousness which remains free and at the same time has an active role. Sāmkhya is not faced with this particular problem since consciousness in that system is not needed to drive the mind-body system, although Sāmkhya cannot explain how moral and spiritual acts take place. This latter problem remains for Vedānta as well, and the solution to this problem is incompatible with a pure observer status for consciousness. Here Vedānta takes refuge in its metaphysical doctrine of ignorance and says that desires do not really exist except through ignorance, the implication being that if desires do not exist then consciousness is not doing anything to control them, it is all just a seeming. This to my mind is a sheer confusion. For whether or not desires exist in the ultimate analysis, they do in a state of ignorance and it is in this state of ignorance that an effort must be made to control them. If this effort is initiated by consciousness then it cannot be described as a pure witness that takes no part in the doings of the empirical self. So in this state an interaction of some sort takes place no matter what prevails in the state of freedom.

The Vedānta doctrine that once an empirical man attains freedom by realising that consciousness is a mere witness he remains free even though he remains empirical to the extent that he continues to have a mind and body seems to me highly implausible. What is not implausible is that a man may come to enjoy within himself a state of awareness that something in him remains unconditioned by empirical limitations and he can do this by virtue of the witnessing function of consciousness which withdraws its identification with these limiting conditions. It remains possible for men to enjoy such freedom to a greater or lesser degree. But it seems, to say the least, highly unlikely that a state of freedom can be attained once and for all by an empirical man. And this for the reason that consciousness in an empirical man is not just a pure witness, it is also involved in the activities of the body and the mind and remains so involved as long as mind and body continue to exist.

It seems to me that the term 'I' as it is used to refer to ourselves can he said to have two distinct components. On the one hand it refers to the empirical ego insofar as it consists of a man's modal consciousness of his so-called states of consciousness which succeed one another, but which do not by themselves suffice, so it seems to me, to supply the sense of identity that a man has when by 'sense of identity' is meant the feeling that a man has of remaining an identical person despite all sorts of changes that happen to him. On the other hand it refers to the function of consciousness in us which is identical as a function of manifestation and to which the contents of a man's experience are present. It is from the presence of this function in man that he gets his sense of identity (in the sense defined above) despite the fact that physical and mental states which give content to a man's personality change. But the empirical man is not an empty pole of identity, he has also a sense of unity as a distinct personality constituted of the diverse contents of his experience all felt to be his own. Nevertheless, unity of personality and identity are somewhat different concepts. Take, for instance, a man who has suffered an accident and as a consequence loses his memory. As this man regains consciousness he cannot connect himself up with his past or any projected future like any normal person, and as a result his sense of being an empirical self, an individual with a personal history, suffers. But he still retains a sense of identity and has no difficulty in referring to himself as 'I' although he cannot say 'I so and so' expressing a certain unity of personal history. He cannot give himself any descriptive characteristics that would particularise him as a distinct empirical

personality and is in the unusual situation of having to accept information about himself from other people if they are forthcoming, but he still takes himself as an identical being whoever that being may be. The unity of his ego as a particular person is lost as the internal states which constitute the elements of his unity are lost and cannot be illumined by consciousness, temporarily at any rate, so long as they remain lost. But his sense of identity remains, despite loss of unity, and this sense of identity comes from the identity of consciousness as a self-manifesting function. If so, a conscious enjoyment of oneself as such a function and the sense of freedom inherent therein, does not seem to be an impossibility and yet man, the empirical man in the totality of his being, is not just this and so his freedom can never be absolute as claimed by those who identify the empirical man with his consciousness.

Thus the empirical man seems to have a paradoxical nature, he is both transcendence and mind-body system functioning according to physico-chemical laws, he is both conditioned and unconditioned. To say that his essence lies in transcendence or in being unconditioned is only to say that he is distinguished from other things and creatures with whom he has a lot in common by having these additional properties, it is not to say that he is transcendence and freedom through and through. However, Vedānta has more reason than Sāmkhya to recommend transcendence and freedom since it believes freedom to be a heightened state of awareness of joy. If it is really such a state then a man who pursues happiness may think that it is worth a try. But if a man does not care for transcendence or freedom of this variety we ought not to say that he is mistaken about his real nature – unless the term real is used as a value term – since it suggests that the limitations inherent in a man's psychophysical system are not a real part of him. All we can say is that a man who does not care to realise that he is free in this ontological sense has failed to realise some possibilities of his nature as a human being, granted of course that transcendence and freedom are facts about human nature.

4
Some Concluding Remarks

We have considered in the foregoing chapters three views of consciousness and freedom, beginning with the materialist view that human consciousness and freedom need not be in any way considered as facts *sui generis* and ending with two Indian views of the idealist variety which claim that they must be so considered. The phenomenological-existentialist views considered here come quite close to the Indian views except for the fact that even when claiming an independent and absolute status for consciousness and freedom Husserl and Sartre do not divest man of all empirical properties – no distinction is made by them between the empirical man and the real or essential man to be recovered or discovered from out of the empirical man – as Indian idealists do. Indian idealists define 'self' as consciousness and freedom and as man is self he is said to be nothing but his consciousness and freedom. Sartre too tends to define man in terms of his consciousness and freedom, but he, perhaps with less logical consistency, goes on to include for certain purposes empirical properties in his concept of man, despite this definition.

Consciousness in the materialist-behaviourist scheme is treated in such a way that it becomes more like the mind in Indian philosophy – an organ of intelligent and discriminative reaction to experience constituted of capacities to think, feel and act as well as of various dispositional properties. The fact that human consciousness is necessarily conscious of itself is hardly noticed in this scheme and when noticed it is quickly played down as just another dispositional property to behave in certain ways, or as just an epiphenomenon having no essential function. This scheme succeeds in its explanations as long as consciousness remains just a capacity to respond in appropriate ways, but it seems to me that as soon as it appears as a self-conscious capacity for reflection either on itself or on its experience of the world, not with a view to responding in any particular fashion in respect to any particular object but solely with a view to understanding for the sake

of it the meaning of an experience which one has already had, behaviourist explanation fails to give any account of it precisely because there is no set or type of behaviour into which it can be translated. Also, if it is the case, as I have argued, that man can, as a self-conscious being, reflectively decide to change his character as hitherto existing – this constitutes his freedom – then consciousness does not remain a mere adjunct as in epiphenomenalism but plays an effective role in the life of man.

When materialists do notice, as Armstrong, for instance, that consciousness is also self-consciousness, they try to show that this too is nothing but a central state in the brain without actually attempting to translate it in behavioural or dispositional terms, which of course cannot be done. I have argued in the first chapter why I think that this project fails and I shall not repeat the argument. Materialists and behaviourists do not find man to be in any essential way different from the rest of the world, though of course enormously more complicated. 'Freedom' and 'self-consciousness' are categories which do not apply to the workings of the rest of the world, so they do not apply to men either – tendencies exhibited in such doctrines as that there is no such thing as experience and that men are completely determined in all their functions. Or if they do apply, then they have to be explained in such a way as requires no break with physical laws in understanding a human personality – tendencies shown in such doctrines as that freedom lies in the ability to do what one likes, and this is perfectly compatible with determinism, and that even though self-consciousness is unique to man, it is perfectly explainable in terms of the complexity and organisation of a brain state which produces it.

With Husserl and Sartre consciousness and freedom appear as facts of their own kind, independent and irreducible, hence man cannot be assimilated to the non-human world. Husserl tries to show that consciousness is needed for a comprehension of meaning involved in the use of concepts, for such use needs that certain *a priori* distinctions have to be intuitively grasped by the intentional function of consciousness. This as presented by Husserl is somewhat ambiguous. For even if it is the case that understanding of the meaning of concepts as occur in a language need certain principles of mental organisation which may in some sense be said to be *a priori* inasmuch as they are innate to the human mind and are not simply built up by generalisation from experience, they may and do function quite unconsciously.[1] Even if the

[1] See N. Chomsky, *Language and Mind* (Harcourt, Brace, 1968).

carrying on of these functions by the mind is intentional this intentionality need not be construed as a self-conscious function. (A man in receiving impressions does not say to himself 'I must put these diverse data together to arrive at a unitary structure,' his mind does the job without his knowing about it.) In other words, meaning of experience can be grasped by the mind unconsciously even when experience involves the use of concepts.

But Husserl also talks about the constitution of an object in cognition which to my mind means 'reconstitution of an object in reflective experience', and this does involve an intuitive grasping of essential structures of things on a self-conscious plane. It is insofar as human beings are capable of this sort of thing and not simply so far as they are capable of using concepts in speech that a transcendental function of consciousness may be conceived to be logically presupposed (meaning that its presence has to be acknowledged as being there as a fact of its own kind). As human beings are capable of reflection, the reality of consciousness cannot be denied.

Sartre rightly points out that all consciousness is self-consciousness even on a non-reflective plane, and as this is intuitively evident there is no need to find this out through a logical operation on man's knowledge performance. Consciousness is not only aware of itself as awareness it is aware of itself as not being the non-conscious (in-itself), and this awareness constitutes the essential characteristic of the being of consciousness. Those who deny this intuitively evident fact are acting on bad faith, according to Sartre. Since the essential nature of consciousness is its transcendence of things which it knows as other than itself, its reality cannot be derived from the thing. And as consciousness is not a thing but a function of transcendence whereby things are revealed a conscious man is necessarily a free man, standing outside the determined chain of the causal series which constitutes the thing world.

Sartre and Husserl do not make any distinction between the (1) self-conscious and (2) non-self-conscious functions of the human mind and they seem to suggest in their doctrine of the primacy of consciousness that even in respect of (2) consciousness plays a controlling role. The behaviourists and materialists however show that a great many functions of mind, involving intelligence, purpose, discrimination, etc., can be explained on neurophysiological principles and that these functions do not need consciousness (meaning self-conscious awareness) as an explanatory hypothesis. Surprisingly enough an Indian school of thought, Sāmkhya, which is idealist insofar as it is concerned in its

philosophising with liberating man from all limiting conditions imposed on him by his physical nature, agrees with materialism in this respect. All that belongs to mind, it says, comes from matter and its principles, and consciousness plays no part whatsoever in the functioning of the psychophysical system. But as one is immediately aware of the fact of consciousness which it will be self-contradictory to deny, there is in the world consciousness as well as a psychophysical system. And as man can say 'I' only insofar as he is conscious, man is his consciousness and not his psychophysical system which he only mistakenly thinks to be himself. This mistake arises because his consciousness is reflected in this system revealing its states, which states then become 'states of consciousness' through a process of misappropriation. However, all that belongs to the empirical man, his physical and mental states and dispositions, remain a function of the material principle, just the same as everything else in the world.

It is interesting to note that Sartre observed that an original ambition of man is to be God, that is, a being who is completely free and who is the source of his own existence. This he declared to be an impossibility. Sāmkhya and Vedānta aim precisely at this. Sāmkhya of course does not believe in God but Vedānta does – God means the Absolute in Vedānta – and it does assert that man in his true essence is identical with the Absolute. The status which both Sāmkhya and Vedānta recommend for man is one in which one's existence is identical with one's essence as consciousness and freedom, as those aspects of man's existence which limit him are to be rejected as forming no part of man's true self. Unlike Sartre they believe that it is possible for man to achieve this state and they declare this achievement to be the true goal of human existence. Sartre does not believe that this is possible for he does not want the empirical man to cease to be empirical. He wants man to work out his paradoxical destiny as best as he can within the human condition, accepting in complete integrity that life is really absurd in the sense of being without cause, without reason and without justification. For all their similarities, and there is a good deal in common between their notions of consciousness and freedom, Sartre would have nothing to do with the Sāmkhya or Vedānta scheme of detachment between man's consciousness and his contingency. No doubt Sartre displays a tendency to play down man's contingency when he defines him as completely self-conscious and completely free. But Sāmkhya and Vedānta carry this definition to its logical extreme and declare that man's nature is pure consciousness and that realisation of freedom lies in nothing short of altogether transcending the human condition. But although Sartre's freedom functions within the human condition and it does not require man to go beyond it, it seems to me that Sartre too makes impossible demands on the empirical man. Sartrean transcendence and authenticity, conditions integral to the concept of remaining in freedom, require man to be continually ahead of himself. And this seems to be equally beyond the human capacity. Indeed Sartre himself laments that most men give up and live in bad faith. Bad faith is such a pervasive condition of human living, according to Sartre, that it seems that his authentic living also requires man to go beyond the human condition in some fashion.

For Sartre and Husserl, intentionality, the direction of consciousness towards an object, even if the object be a non-existent something, is an integral aspect of it although they both speak of the self-manifesting function of consciousness as well. For Sāmkhya and Vedānta intentionality only describes a characteristic of what they call 'modal consciousness' and it is in actual fact a feature of intelligence which is a directional function. Pure consciousness is a state of self-manifestation which is a witnessing function that need not witness anything but itself. That it does witness objects is only a contingent but not a logically necessary fact about consciousness the essence of which lies in self-manifestation. It is compared to a light which is turned on an empty stage wherein it reveals nothing but itself. Their emphasis on this aspect of consciousness to the exclusion of the intentional function arises out of their anxiety to secure a state of complete freedom for man, which cannot be attained if the empirical world as the intentional object of consciousness is integrated into its very being.

We can thus see that there are certain ways of thinking about the mind, both materialist and non-materialist, which turn up in very different intellectual and cultural contexts, although the uses to which these thoughts are put in one's view of man and his possibilities tend to be different according to one's cultural presuppositions and valuational preferences which may be purely personal. It is these presuppositions and preferences which condition one's decision as to the purpose for which a study of the mind is to be undertaken and it is in relation to this purpose that one picks up the relevant dimension of mind to be studied, the other dimension being irrelevant and so of no interest or consequence.

The material dealt with in this book shows three basic purposes and themes controlling one's approach to the problem of mind: (1) The

behaviourist-materialist theme is that man is completely continuous with the rest of the things in the world and that he operates exactly on the same principles. This can be demonstrated by studying human behaviour in terms of scientific methodology. This approach to mind represents an enormous gain in our understanding of the mind because it enables us to have experimentally verified knowledge in place of vague intuition and introspection which people who believe in the privacy of mind apply to the study of mind to no visible effect. (2) The phenomenological-existentialist theme is that man is discontinuous with the thing world and an understanding of what he is like needs a basically different approach. Man intuitively grasps himself as different from the world of things and it is through his intentional and intuitive capacities that a shared world of experience comes to have the kind of significance that this world has for man. This can be demonstrated by studying consciousness and its functions according to the phenomenological method, if it needs any demonstration, for the peculiarity of consciousness is that it is present to itself which fact is intuitively grasped. (3) The Indian theme is that man is not the limited and conditioned creature that he takes himself to be. His limitations come from the limitations of the determined system that the mind-body part of his personality represents. But this can be overcome, for man's essence lies in consciousness and consciousness can be conscious of itself as just a witness. It is by identifying oneself with the witnessing function of consciousness that one can overcome the limitations of one's physical being and in the process come to realise oneself as unlimited and unconditioned as all of existence instead of as a limited being subject to the causal functioning of the physical aspect of one's personality.

But what is significant is that the materialist dimension of mind can come to be recognised even in a context of inquiry the presuppositions of which are transcendental in character, equally, that the intuitively grasped self-transcending character of consciousness can be recognised in a context where man is not expected to function as anything other than an empirical being who in spite of being conscious and free is limited by facticity and contingency. So both these dimensions of mind are recognisable facts and theories of mind which stress one of these are often based on a decision to disregard the other. The controlling theme behind one's study of man decides what would be considered important or sometimes even what would be noticed, but there is no one-to-one correspondence between one's cultural conditioning and one's capacity to notice things. The materiality of mind can come to be noticed in an intellectual climate which is predominantly idealist and transcendence of consciousness in a cultural context which is predominantly materialist.

Of course even if Sāmkhya recognises that the mind is material its real concern is not to preach that but to say that man is not mind but consciousness. But as already remarked this is a value judgment about what man ought to be like and not a statement about what he is like and certain facts about how a man actually functions must have been noticed for a philosopher to come to the decision that the mind is material. Similarly in a rationalist-intellectual climate a philosopher is most unlikely to come to the conclusion that man ought to try and transcend his material condition altogether, but this need not prevent him from noticing that man is a self-conscious being possessed of a capacity for transcendence and freedom.

Once the valuational character of theories of mind is understood there need not be any rigid either/or about the dimension of mind one ascribes to man, for it is possible for one to recognise that the other dimension is there and yet find that it has no importance or interest for one's particular purposes. There is no *a priori* reason why a certain dimension of mind cannot be material while another is conscious, capable of transcendence and freedom, understood not only in an empirical context but also in a transcendental one. And to say that man is both a material system and a conscious being capable of transcendence and freedom is not only to claim that these are facts about man, but also to express a valuational attitude that both these facts are interesting and important for one's understanding of man.